 66–10,177

JACOBS, Marc Quillen, 1938–
THE EXISTENCE OF OPTIMAL CONTROLS.

The University of Oklahoma, Ph.D., 1966
Mathematics

University Microfilms, Inc., Ann Arbor, Michigan

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE EXISTENCE OF OPTIMAL CONTROLS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

MARC QUILLEN JACOBS

Norman, Oklahoma

1966

THE EXISTENCE OF OPTIMAL CONTROLS

APPROVED BY

George M. Ewing

[illegible]

Allen S Davis

Kenneth R. Merrill

Miriam C. [illegible]

John [illegible]

DISSERTATION COMMITTEE

ACKNOWLEDGEMENTS

I wish to express my deep appreciation to Dr. G. M. Ewing, who directed my research for this dissertation, for his patient encouragement and the many useful suggestions which he offered during our stimulating weekly conferences while this dissertation was being prepared. My thanks also go to the sponsors (Dr. G. M. Ewing and Dr. W. T. Reid) of the weekly seminar in Differential Equations and Variational Theory held at the University of Oklahoma from September 1964 to August 1965, for the opportunity to present many of the results obtained in this work. Dr. Ewing, Dr. Reid, and the other participants in this seminar are to receive my special thanks for the critical appraisal which they gave to my presentations.

This work was partially sponsored by the Office of Aerospace Research of the Air Force Office of Scientific Research, the United States Air Force, under Grant AF-AFOSR-211-63, during the period from June 1963 through August 1964, and by a National Aeronautics and Space Administration Fellowship, during the period from September 1964 to September 1965. I am very grateful to these governmental agencies for the aid which I received from them.

TABLE OF CONTENTS

THE EXISTENCE OF OPTIMAL CONTROLS

CHAPTER I

INTRODUCTION

It is very easy to form an intuitive notion of what is meant by an optimal control problem. Man frequently tries to attain his desired goals in the "best" possible manner. Often such attempts are futile because of the complexity of the problem or because there does not exist any "best" way of attaining the desired goal. However, in a great many of the physical processes occurring in technology, there is more than one policy (or control) which will enable the system in question to attain a certain desired objective. Among the "available" policies with which the objective can be attained an "optimal policy" is sought, i.e., a policy in the collection of all "available" policies which will enable the controlled system to attain the desired objective, which minimizes the "cost" of using such a policy or maximizes some other given "criterion of optimality." Here we must understand the term "cost" in a very broad sense. It is evident that in order to bring this discussion within the purview of mathematics we must give precise definitions of what is meant by such terms as "cost," "best," and "available policies." Moreover, a precise mathematical model must be given to describe the <u>controlled system</u> (the system or plant with which one attempts to reach the desired goal, target, or objective). We shall use such terms as <u>admissible controls</u>, <u>available policies</u>, and admissible steering functions interchangeably, and, of course, these classes of functions must be defined in mathematical terms. It will come as no surprise to the reader that the problem of finding an "accurate" mathematical model for a given physical process often is one of considerable difficulty. We shall not concern ourselves in this work with the question of whether our mathematical models accurately portray certain physical processes, but rather we shall make our work much easier by accepting certain mathematical models as given and our studies will begin at this point. In other words the problems which we will consider in this dissertation are of a purely mathematical nature. We have not of course chosen them with caprice, and there is actually good reason to believe that they are fairly realistic idealizations of a variety of material systems.

Basically only two types of optimal control problems will be discussed in this work. We shall refer to these two problems as the fundamental problem of optimal control, and the fundamental problem of optimal control in a system with delayed argument. First we shall give, in broad outline, the formulation of the fundamental problem of optimal control. Our formulation differs in several details from that given by Pontryagin* and his collaborators [XV]. We assume the control system can be described by a system of real ordinary differential equations of the form

$$\mathring{x} = f(t,x,u) \tag{1.1}$$

where $x = (x^1,x^2,\ldots,x^n)\varepsilon R^n$, $u = (u^1,u^2,\ldots,u^m)\varepsilon R^m$, and f is the mapping, $f = (f^1,f^2,\ldots,f^n)$: $R\times R^n\times R^m \longrightarrow R^n$, and $t\varepsilon R$. Of course, $\mathring{x}$

*All references are to the bibliography at the end of this dissertation. The symbolism ["Roman numeral"] will refer to the book list in the bibliography, and ["Arabic numeral"] will refer to the list of articles in the bibliography.

means differentiation with respect to the independent variable t. The symbol u denotes the <u>control parameter</u>. Let $\Lambda = [T_0, T_1] \subset R$ be a fixed compact interval, let $\mathcal{C}(R^q)$ denote the collection of closed subsets of R^q, and suppose there is given a mapping $\Omega: \Lambda \longrightarrow \mathcal{C}(R^m)$, $\Omega(t) \neq \phi$ for each $t \epsilon \Lambda$. Let Γ be a closed subset of $\Lambda \times \Lambda \subset R^2$ such that $(t_0, t_1) \epsilon \Gamma$ implies $t_0 \leq t_1$. Now we define $\mathcal{U}(\Omega, \Gamma)$, to be the set

$$\{(u,(t_0,t_1)) \mid u: [t_0,t_1] \longrightarrow R^m \ \underline{\text{is bounded and measurable}},\ (t_0,t_1)\epsilon\Gamma,\ u(t)\epsilon\Omega(t),\ \underline{\text{when}}\ t_0 \leq t \leq t_1\}.$$

Let $x_0 \epsilon R^n$, and we assume (1.1) has the property

(1.2) $\forall\ (u,(t_0,t_1)) \epsilon\ \mathcal{U}(\Omega,\Gamma)$, $\exists$ a unique absolutely continuous <u>response</u> $x(\cdot, u)$ satisfying (1.1) almost everywhere (a.e.) on the domain $[t_0, t_1]$, and satisfying the initial condition

$$(1.3) \qquad x(t_0, u) = x_0.$$

Suppose there is given also a "target mapping" $F: \Lambda \longrightarrow \mathcal{C}(R^n)$; then the set of admissible controls $\hat{\mathcal{U}}(\Omega,\Gamma)$ is defined to be that particular subset of $\mathcal{U}(\Omega,\Gamma)$ satisfying the additional property that

$$(1.4) \qquad (u,(t_0 t_1)) \epsilon\ \hat{\mathcal{U}}(\Omega,\Gamma) \Longrightarrow x(t_1,u) \epsilon F(t_1).$$

The term "target mapping" is somewhat suggestive in that among the motivating raw problems for control theory is the question of optimal trajectories for an anti-missile missile. Hence $F(t)$ might be a subset of state space, associated with but generally larger than the actual target such that $x(t_1,u)\varepsilon F(t_1)$ implies success for the mission.

Suppose there is given a mapping $K: \mathcal{U}(\Omega,\Gamma) \longrightarrow R$ which will be called the cost function (or criterion of optimality) associated with the control system (1.1). Three important and interrelated questions arise:

(1.5) (i) Is $\hat{\mathcal{U}}(\Omega,\Gamma)$ empty?

(ii) Does there exist an optimal control in $\hat{\mathcal{U}}(\Omega,\Gamma)$, i.e., is there a $(u,(t_0,t_1))\varepsilon\ \hat{\mathcal{U}}(\Omega,\Gamma)$ such that

$$K(u,(t_0,t_1)) \le K(v,(t_0^*,t_1^*)),\quad \forall (v,(t_0^*,t_1^*))\varepsilon\ \hat{\mathcal{U}}(\Omega,\Gamma)$$

(global minimum), or is there a $(u,(t_0,t_1))\varepsilon\ \hat{\mathcal{U}}(\Omega,\Gamma)$ such that

$$K(u,(t_0,t_1)) \ge K(v,(t_0^*,t_1^*)),\quad \forall (v,(t_0^*,t_1^*))\varepsilon\ \hat{\mathcal{U}}(\Omega,\Gamma)$$

(global maximum)?

(iii) How may an optimal control be characterized when it exists?

The fundamental problem(s) of optimal control is (are)

(1.6) $K(u,(t_0,t_1))$ = global minimum (or maximum) on $\hat{\mathcal{U}}(\Omega,\Gamma)$.

In this dissertation, as the title indicates, we shall focus attention on (1.5) (ii); the important related questions (1.5) (i) and (1.5) (iii) will be largely ignored. The main purpose of this work is to discover how additional hypotheses may be made so that an affirmative answer to (1.5) (ii) can be given.

We shall also consider control systems which are complicated by the effect of a "delayed" argument, i.e., control systems which can be described by a system of real ordinary differential-difference equations of the following type

$$(1.7) \qquad \mathring{x}(t) = f(t, x(t), x(t - \omega), u(t)), \qquad \omega > 0,$$

and $x = (x^1, x^2, \ldots, x^n) \varepsilon R^n$, $u = (u^1, u^2, \ldots u^m) \varepsilon R^m$, $f = (f^1, f^2, \ldots f^n) : R \times R^n \times R^n \times R^m \longrightarrow R^n$, and $t \varepsilon R$. If $(u, (t_0, t_1)) \varepsilon \mathcal{U}(\Omega, \Gamma)$ and there is given a suitable <u>initial function</u> $\phi : [t_0 - \omega, t_0] \longrightarrow R^n$, then we assume there is a unique response $x(\cdot, \phi, u)$, which is absolutely continuous on $[t_0, t_1]$, and such that

(1.8) (i) $x(t, \phi, u) = \phi(t)$, $t_0 - \omega \leq t \leq t_0$

(ii) $x(\cdot, \phi, u)$ satisfies (1.7) a.e. on $[t_0, t_1]$.

We define $\mathcal{U}^*_\phi(\Omega, \Gamma)$ to be that particular subset of $\mathcal{U}(\Omega, \Gamma)$ such that

$$(1.9) \qquad (u, (t_0, t_1)) \varepsilon \mathcal{U}^*_\phi(\Omega, \Gamma) \Longrightarrow x(t_1, \phi, u) \varepsilon F(t_1).$$

We shall also permit ϕ (the initial function) to vary over a certain prescribed function space Φ. Thus we suppose there is given a mapping $K^* : \mathcal{U}^*_\phi(\Omega, \Gamma) \longrightarrow R$; then <u>the fundamental problem(s) of optimal control in a system with delayed argument is</u> (are)

(1.10) $K^*(u,(t_0,t_1))$ = <u>global</u> <u>minimum</u> (<u>or</u> <u>maximum</u>) on $\bigcup_{\phi\varepsilon\Phi}\mathcal{U}_\phi^*(\Omega,\Gamma)$.

Again we shall only be concerned with developing hypotheses sufficient to guarantee that there is a $(u,(t_0,t_1))\varepsilon\bigcup_{\phi\varepsilon\Phi}\mathcal{U}_\phi^*(\Omega,\Gamma)$ satisfying (1.10) (minimum or maximum).

Our purpose in Chapter II is twofold: First we develop a method of studying the convergence of the mappings Ω, F, and then we compare our methods with other existing methods. Second the fundamental definitions and compactness theorems (for certain function spaces) which will play an essential role in the developments of the subsequence chapters are given. In Chapter III we shall be concerned exclusively with existence theorems for problem (1.6) when the corresponding control system is <u>linear</u>. Chapter IV deals with the existence theory of the fundamental problem of optimal control when the control system (1.1) is permitted to be non-linear.

With a few exceptions [13, 14] very little has been done in the literature concerning existence theory for problem (1.10) in which the control system is given by a system of differential-difference equations. We feel that it is therefore entirely appropriate to include some existence theorems for problem (1.10) in this work. This is the subject considered in Chapter V. After reading Chapter V the reader will see clearly how many of the existence theorems concerned only with the fundamental problem of optimal control can be extended to cover the fundamental problem of optimal control in a system with delayed argument.

We shall point out in the chapters in which our theorems appear (and are proved) how our existence theorems differ from (or are related to) the known results obtained by others. The hypotheses for the theorems developed here are so elaborate that to relate our work in any greater detail to the other results in this field is not, at this juncture, economical.

CHAPTER II

DEFINITIONS AND FUNDAMENTAL THEOREMS

In this chapter we shall give the definitions of the fundamental terms which will be used repeatedly. In addition several theorems will be proved in this chapter which are essential to the remainder of this work. Of special importance is the section dealing with upper and lower semicontinuity in the lattice of closed subsets of a given topological space. Also since our primary purpose is to prove existence theorems for optimal control problems, the reader can expect that compactness of certain function spaces will play an important role. We give in this chapter the proof of one such compactness theorem (Theorem 2.2.3) which we shall require in the proof of several existence theorems. Theorem 2.2.4 also will play a significant part in our future developments.

2.1 Upper and Lower Semicontinuity in the Lattice of Closed Subsets of a Topological Space.

It will be convenient to draw upon the theory of uniform spaces (all nomenclature in Section 2.1 referring to uniform spaces is the same as in Bourbaki [V, Chapter II]) in order to discuss two different types of upper (lower) semicontinuous mappings from a given topological space to the complete lattice of closed subsets of another topological space. Our applications in Chapters III-V are limited to a discussion of the convergence of nonempty classes of closed subsets of euclidean space of n-dimensions; however, it is felt that the introduction of uniform structures at this point serves to simplify much of our later work as well as to illuminate the relationships between various types of convergence in the set of all closed subsets of euclidean space of n-dimensions. Uniform structures enter our study in a natural way whenever we begin to discuss the continuity or upper semicontinuity with respect to inclusion (see the definitions given below) of the "motion" of certain one-parameter families of subsets of a given uniform space. Such situations are common in the formulation of optimal control problems [7, 11, and 13]. The concept of upper semicontinuity with respect to inclusion permeates the literature in the mathematical theory of optimal control. The term suggests that what is meant is the usual concept of upper semicontinuity of a function F from a topological space X to the complete lattice of closed subsets (ordered by inclusion) of a given topological space Y [XIV, pg. 73 ff.]. This turns out not to be the case, but we have discovered that what we shall need in this work is not the so-called upper semicontinuity with respect to inclusion, but rather upper semicontinuity as it is normally defined for a function F on a topological space X to a complete lattice L [XIV, pg. 73 ff.]. We do not, however, wish to make an issue out of which concept should be called upper semicontinuity, since there appear to be ample reasons to support using this terminology to denote either concept. We shall also show that, in general, upper semicontinuity with respect to inclusion is a strictly stronger property than upper semicontinuity in the usual sense.

At this point some introductory terminology is in order. If $\mathcal{B}$ is a filter base on a set E, then we use $[\mathcal{B}]$ to denote the smallest filter on E containing $\mathcal{B}$ (see [V, Chapter I] for any terminology related to filters).

Denote by Δ the "diagonal of $E\times E$," i.e.,

$$\Delta = \{ (x, x) \mid x \in E \}.$$

We then recall the definition of a uniform structure on a set E. A filter $\mathcal{S}$ on $E\times E$ is a uniform structure on E iff

$$\begin{array}{ll} \text{(U1)} & [\Delta] \supset \mathcal{S} ; \\ \text{(U2)} & \mathcal{S}\circ\mathcal{S} = \mathcal{S} ; \\ \text{(U3)} & \mathcal{S} = \mathcal{S}^{-1} , \end{array}$$

where

$$\mathcal{S}^{-1} = \{ V^{-1} \mid V \in \mathcal{S} \}$$
$$\mathcal{S}\circ\mathcal{S} = [\{ V\circ W \mid V, W \in \mathcal{S} \}]$$

and for $V, W \subset E\times E$

$$V\circ W = \{ (x, y) \mid \exists\, z \in E \ni (x, z) \in V \text{ and } (z, y) \in W \} ,$$
$$V^{-1} = \{ (y, x) \mid (x, y) \in V \} .$$

If $V \subset E\times E$, $A \subset E$, it is convenient to define

$$V[A] = \{ y \mid y \in E \ \exists\, x \in A, (x, y) \in V \} .$$

Using this notation a uniform structure $\mathcal{J}$ on E induces a topology on E in the usual way, viz., by defining the neighborhood filter at $x \in E$, [V, Chapter I, pg. 3], to be

$$\mathcal{J}[x] = \{ V[x] \mid V \in \mathcal{J} \}.$$

The topology (collection of open sets) associated with this neighborhood filter will be denoted by $\tau_{\mathcal{J}}$.

DEFINITION 2.1.1. For any topological space (E, τ) we define

$$\mathcal{C}(E) = \{ A \mid A \subset E,\ A \ \underline{\text{is}}\ \tau\text{-}\underline{\text{closed}} \}.$$

Thus if $(E, \mathcal{J})$ is a uniform space, we shall show how a natural uniform structure may be defined on $\mathcal{C}(E)$. The uniform structure which we have in mind is discussed by Bourbaki [V, Chapter II, pg. 97], Michael [12, pg. 155 ff.], and the Robertsons [16, pg. 322], so we will not exhibit many details. Define for $J \in \mathcal{J}$

$$\mathbb{V}_J = \{ (A, B) \mid A, B \in \mathcal{C}(E),\ J[A] \supset B \ \&\ J[B] \supset A \}.$$

Furthermore let $\mathcal{B}(\mathcal{J})$ be defined to be the set $\{ \mathbb{V}_J \mid J \in \mathcal{J} \}$. Then it can easily be shown that $[\mathcal{B}(\mathcal{J})]$ is a uniform structure on $\mathcal{C}(E)$. We shall adopt a notation suggested by Michael [12] and define

$$2^{\mathcal{J}} = [\mathcal{B}(\mathcal{J})]. \tag{2.1.2}$$

Then the system of neighborhoods at $A \in \mathcal{C}(E)$ (the neighborhood filter at $A \in \mathcal{C}(E)$) is given by

$$2^{\mathcal{J}}[A] = \{ \mathbb{H}[A] \mid \mathbb{H} \in 2^{\mathcal{J}} \}. \tag{2.1.3}$$

Next we assume (E,τ) is an arbitrary topological space. Then $\mathcal{G}(E)$ can be made into a complete lattice in the usual way, viz., by defining for $A,B\in\mathcal{G}(E)$

$$A\leq B \text{ iff } A\subset B.$$

Thus if $F_i\in\mathcal{G}(E)$, $i\in I$ = an index set, then

$$\inf\{F_i \mid i\in I\}= \bigcap_{i\in I}F_i \tag{2.1.4}$$

and

$$\sup\{F_i \mid i\in I\} = \overline{\bigcup_{i\in I}F_i}^{\tau} \tag{2.1.5}$$

where the symbol $\overline{A}^{\tau}$ denotes the τ-closure of $A\subset E$.

DEFINITION 2.1.2. Let (X,τ_1), (E,τ_2) be topological spaces. Suppose F is a mapping, $F:X\to\mathcal{G}(E)$. Let $\mathcal{V}_1(a)$ denote the τ_1-neighborhood filter at $a\in X$. Then for $a\in X$

$$\limsup_{x\to a} F(x) = \inf\{\sup\{F(x)\mid x\in G\}\mid G\in\mathcal{V}_1(a)\}.$$

In a similar manner we define

$$\liminf_{x\to a} F(x) = \sup\{\inf\{F(x)\mid x\in G\}\mid G\in\mathcal{V}_1(a)\}.$$

DEFINITION 2.1.3. Using the terminology of Definition 2.1.2 we say (by an <u>abus de langage</u>, cf. the comment after Example 2.1.2 which appears later in this chapter) that a mapping F from a topological space X to the collection of closed subsets, $\mathcal{G}(E)$, of another topological space, E, is <u>upper semicontinuous</u> at $a\varepsilon X$ (abbreviated F is usc at $a\varepsilon X$) iff

$$\limsup_{x\to a} F(x)\leqslant F(a).$$

Similarly F is <u>lower semicontinuous</u> at $a\varepsilon X$ (abbreviated F is lsc at $a\varepsilon X$) iff

$$\liminf_{x\to a} F(x)\geqslant F(a).$$

Our definitions of lsc and usc are precisely those given by McShane and Botts [XIV, pg. 73].

DEFINITION 2.1.4. Let (X,τ) be a topological space, $(E,\mathcal{J})$ be a uniform space, and let F be a mapping, $F:X\to\mathcal{C}(E)$. Suppose ${}^{\tau}\mathcal{V}(a)$ denotes the τ-neighborhood filter at $a\varepsilon X$. We shall say F is <u>upper semicontinuous with respect to inclusion</u> at $a\varepsilon X$ (abbreviated F is usci at $a\varepsilon X$) iff

$$\forall J\varepsilon\mathcal{J}\ \exists G\varepsilon{}^{\tau}\mathcal{V}(a)\ni x\varepsilon G\Rightarrow J[F(a)]\supset F(x).$$

Similarly we shall say F is <u>lower semicontinuous with respect to inclusion</u> at $a\varepsilon X$ (abbreviated F is lsci at $a\varepsilon X$) iff

$$\forall J\varepsilon\mathcal{J}\ \exists G\varepsilon{}^{\tau}\mathcal{V}(a)\ni x\varepsilon G\Rightarrow J[F(x)]\supset F(a).$$

We introduce an additional notational device. Let $G \subset X$; then we define

$$F(G) = \bigcup_{x\in G} F(x) \subset E,$$

and

$$\{F(G)\} = \{F(x) \mid x\in G\} \subset \mathcal{E}(E).$$

DEFINITION 2.1.5. $F(A)$ (= the collection of all filters on the set A) may be partially ordered by a relation, $\leq$, defined by

$$\mathcal{F} \leq \mathcal{G} \text{ iff } \mathcal{F} \subset \mathcal{G},\ \mathcal{F}, \mathcal{G} \in F(A).$$

Using this partial ordering on $F(\mathcal{E}(E))$ we make the following definition

DEFINITION 2.1.6. Let (X,τ_1), (Y,τ_2) be two topological spaces, and let F be a mapping, $F:X \to Y$. Suppose $\mathcal{V}_1(x), \mathcal{V}_2(y)$ denote the neighborhood filters at $x\in X$, $y\in Y$ respectively. Then F is <u>continuous</u> at $a\in X$ iff

$$F(\mathcal{V}_1(a)) \geq \mathcal{V}_2(F(a))$$

where $F(\mathcal{V}_1(a)) = [\{F(G) \mid G\in \mathcal{V}_1(a)\}]$.

There are several other concepts of upper and lower semicontinuity of mappings $F:X \to \mathcal{E}(E)$, e.g., see [4, pg. 67 ff., 12, pg. 179, 9 pg. 148], but as we pointed out earlier the definitions which we have given seem to be more suited to our needs.

THEOREM 2.1.1. *Let* (X,τ_1) (E,τ_2) *be topological spaces. Let* $\mathcal{B}(a)$ *be a filter base to the* τ_1*-neighborhood filter,* $\mathcal{V}_1(a)$, *at* $a\varepsilon X$. *Then* $F:X\to\mathcal{C}(E)$ *is usc at* $a\varepsilon X$ *iff*

$$\bigcap_{B\varepsilon\mathcal{B}(a)}\overline{F(B)}^{\tau_2}\subset F(a).$$

Also F *is lsc at* $a\varepsilon X$ *iff*

$$\overline{\bigcup_{B\varepsilon\mathcal{B}(a)}\bigcap_{x\varepsilon B}F(x)}^{\tau_2}\supset F(a).$$

Proof: We first establish the validity of the equalities

$$\text{(2.1.6)}\quad \text{(i)}\quad \bigcap_{G\varepsilon\mathcal{V}_1(a)}\overline{F(G)}^{\tau_2}=\bigcap_{B\varepsilon\mathcal{B}(a)}\overline{F(B)}^{\tau_2};$$

$$\text{(ii)}\quad \overline{\bigcup_{G\varepsilon\mathcal{V}_1(a)}\bigcap_{x\varepsilon G}F(x)}^{\tau_2}=\overline{\bigcup_{B\varepsilon\mathcal{B}(a)}\bigcap_{x\varepsilon B}F(x)}^{\tau_2},$$

from which the two assertions of the theorem follow immediately. Since $\mathcal{B}(a)$ is a filter base for $\mathcal{V}_1(a)$, it follows that $\mathcal{B}(a)\subset\mathcal{V}_1(a)$. Therefore

$$\bigcap_{B\varepsilon\mathcal{B}(a)}\overline{F(B)}^{\tau_2}\supset\bigcap_{G\varepsilon\mathcal{V}_1(a)}\overline{F(G)}^{\tau_2}$$

and

$$\overline{\bigcup_{G\varepsilon\mathcal{V}_1(a)}\bigcap_{x\varepsilon G}F(x)}^{\tau_2}\supset\overline{\bigcup_{B\varepsilon\mathcal{B}(a)}\bigcap_{x\varepsilon B}F(x)}^{\tau_2}.$$

But by hypothesis $[\mathcal{B}(a)]=\mathcal{V}_1(a)$, so that given $G\varepsilon\mathcal{V}_1(a)$ there is a $B\varepsilon\mathcal{B}(a)$ such that $G\supset B$. Thus

$$\overline{F(G)}^{\tau_2}\supset\overline{F(B)}^{\tau_2}$$

and

$$\bigcap_{x\varepsilon G} F(x) \subset \bigcap_{x\varepsilon B} F(x) .$$

The validity of equalities (2.1.6) (i) and (ii) now follows.

Theorem 2.1.2 below suggests what appears to have been the motivation for Definition 2.1.4. [cf. 7, 11, and 15].

THEOREM 2.1.2. <u>Let</u> (X,τ) <u>be a topological space, let</u> $(E,\mathcal{J})$ <u>be a uniform space, and let</u> F <u>be a mapping,</u> $F:X \to \mathcal{C}(E)$. <u>Then</u> F <u>is continuous at</u> $a\varepsilon X$ <u>iff</u> F <u>is usci and lsci at</u> $a\varepsilon X$.

Proof: As usual we let $\mathcal{V}(a)$ denote the τ-neighborhood filter at $a\varepsilon X$. Then according to Definition 2.1.6 and relation (2.1.3) F is continuous at $a\varepsilon X$ means

$$F(\mathcal{V}(a)) \geqslant 2^{\mathcal{J}} [F(a)].$$

Consequently given $J\varepsilon \mathcal{J}$, $\exists\, G\varepsilon\mathcal{V}(a) \ni \{F(G)\} \subset \mathcal{W}_J[F(a)]$. Therefore $x\varepsilon G \Rightarrow J[F(a)] \supset F(x)$ <u>and</u> $J[F(x)] \supset F(a)$. Whence it follows that F is usci and lsci at $a\varepsilon X$. Conversely suppose F is usci and lsci at $a\varepsilon X$. Then given $J\varepsilon \mathcal{J}$ $\exists\, G_1, G_2\varepsilon\mathcal{V}(a) \ni x\varepsilon G_1 \Rightarrow J[F(x)] \supset F(a)$; $x\varepsilon G_2 \Rightarrow J[F(a)] \supset F(x)$. But $G_1 \cap G_2 \varepsilon \mathcal{V}(a)$, and consequently $x\varepsilon G_1 \cap G_2 \Rightarrow J[F(x)] \supset F(a)$ <u>and</u> $J[F(a)] \supset F(x)$. It easily follows now that

$$\mathcal{W}_J[F(a)] \supset \{F(G)\}$$

and therefore that

$$F(\mathcal{V}(a)) \geq 2^{\mathcal{V}}[F(a)] ,$$

which shows that F is continuous at $a\varepsilon X$.

The next theorem shows that the upper semicontinuity of F on X is precisely that property which we shall require in most of our investigations in control theory. Before we state the theorem we remind the reader that a topological space (E,τ) is regular iff for each $a\varepsilon X$ the neighborhood filter, $\mathcal{V}(a)$, at $a\varepsilon X$ can be generated by a filter base of τ-closed subsets of X.

THEOREM 2.1.3. <u>Let</u> (X,τ_1) <u>be a compact regular topological space, and let</u> (E,τ_2) <u>be a topological space. Suppose</u> F <u>is a mapping</u> $F:X\to\mathcal{C}(E)$. <u>Then the following two statements are equivalent:</u>

(i) F <u>is usc on</u> X;

(ii) $\{x_n, n\varepsilon D\}$, $\{P_n, n\varepsilon D\}$ <u>nets in</u> X <u>and</u> E <u>respectively,</u>

$x_n \xrightarrow[\tau_1]{} x_o$, $P_n \xrightarrow[\tau_2]{} P_o$, $P_n\varepsilon F(x_n)$, $n\varepsilon D \Rightarrow P_o\varepsilon F(x_o)$.

Remark: All nomenclature and theorems concerning <u>nets</u> to be used in this work can be found in Kelley [IX, Chapter 2]. The definition of a <u>net</u> is to be found in [IX, pg. 65].

<u>Proof of Theorem 2.1.3</u>: (ii)$\Rightarrow$(i). We show that the under hypothesis (ii) F must be usc at x_o. $\mathcal{V}_1(x_o)$ denotes the τ_1-neighborhood filter at $x_o\varepsilon X$, and we must demonstrate that

$$\bigcap_{G\varepsilon\mathcal{V}_1(x_o)} \overline{F(G)}^{\tau_2} = F(x_o).$$

Since (X,τ_1) is regular $\mathcal{V}_1(x_0)$ has a closed filter basis, say $\overline{\mathcal{B}}$. Then it will suffice (by Theorem 2.1.1) to show

$$\bigcap_{G\varepsilon\overline{\mathcal{B}}} \overline{F(G)}^{\tau_2} \subset F(x_0).$$

Thus let $G\varepsilon\overline{\mathcal{B}}$, $P_0\varepsilon\overline{F(G)}^{\tau_2}$. Then there is a net $\{P_n^G, n\varepsilon D\}$ in $F(G)$ such that $P_n^G \xrightarrow[\tau_2]{} P_0$. $P_n^G\varepsilon F(G), \forall n\varepsilon D \Longrightarrow \exists$ a net $\{x_n^G, n\varepsilon D\}$ (with the same directed set D) such that $P_n^G\varepsilon F(x_n^G)$, $x_n^G\varepsilon G$, $n\varepsilon D$. But G is τ_1-closed, $G\subset X$, and X is compact. Therefore G must be compact [IX, pg. 140]. Therefore there is a subnet of $\{x_n^G, n\varepsilon D\}$, say $\{x_{n_\alpha}^G, \alpha\varepsilon A\}$ such that $x_{n_\alpha}^G \xrightarrow[\tau_1]{} x^G\varepsilon G$.

But then we form the corresponding subnet $\{P_{n_\alpha}^G, \alpha\varepsilon A\}$ of $\{P_n^G, n\varepsilon D\}$ and we must have $P_{n_\alpha}^G \xrightarrow[\tau_2]{} P_0$. Moreover $P_{n_\alpha}^G\varepsilon F(x_{n_\alpha}^G), \forall\alpha\varepsilon A \Rightarrow P_0\varepsilon F(x^G)$ by hypothesis (ii). Now $\{x^G, G\varepsilon\overline{\mathcal{B}}\}$ is a net in X with the property that $x^G\varepsilon G, \forall G\varepsilon\overline{\mathcal{B}}$ (of course, $\overline{\mathcal{B}}$ is directed by $G_1 \leq G_2$ iff $G_2\subset G_1$, $G_1, G_2\varepsilon\overline{\mathcal{B}}$). Therefore $P_0\varepsilon F(x^G)$ for each $G\varepsilon\overline{\mathcal{B}}$, $x^G \xrightarrow[\tau_1]{} x_0$, P_0 = constant net $\xrightarrow[\tau_2]{} P_0$ which gives, in view of (ii), that $P_0\varepsilon F(x_0)$.

Whence

$$\bigcap_{G\varepsilon\overline{\mathcal{B}}} \overline{F(G)}^{\tau_2} \subset F(x_0).$$

(i)$\Rightarrow$(ii). Select $x_0\varepsilon X$, $P_0\varepsilon E$, and nets $\{x_n, n\varepsilon D\}$, $\{P_n, n\varepsilon D\}$ in X and E respectively such that

$$x_n \xrightarrow[\tau_1]{} x_o,\ P_n \xrightarrow[\tau_2]{} P_o,\ P_n \varepsilon F(x_n),\ \forall n\varepsilon D.$$

Now we have by hypothesis (i)

$$\limsup_{x \to x_o} F(x) \leq F(x_o) .$$

Whence

$$\bigcap_{G\varepsilon \mathcal{V}_1(x_0)} \overline{F(G)}^{\tau_2} \subset F(x_o).$$

Let $G\varepsilon\mathcal{V}_1(x_o)$; then there is an $n_o\varepsilon D$ such that $n \geq n_o \Rightarrow x_n \varepsilon G$. Let $S_o = \{n | n \geq n_o,\ n\varepsilon D\}$; then S_o is a directed set and $\{x_n,\ n\varepsilon S_o\}$ is a subnet of $\{x_n,\ n\varepsilon D\}$ and $\{P_n,\ n\varepsilon S_o\}$ is a subnet of $\{P_n,\ n\varepsilon D\}$. Thus

$$\{P_n,\ n\varepsilon S_o\} \xrightarrow[\tau_2]{} P_o,\quad \{x_n,\ n\varepsilon S_o\} \xrightarrow[\tau_1]{} x_o,$$

and $x_n \varepsilon G$, $n\varepsilon S_o$, $P_n \varepsilon F(x_n) \subset F(G)$, $n\varepsilon S_o$. Therefore $P_o \varepsilon \overline{F(G)}^{\tau_2}$. Since $G\varepsilon\mathcal{V}_1(x_o)$ was arbitrary we have

$$P_o \varepsilon \overline{F(G)}^{\tau_2},\ \forall G\varepsilon \mathcal{V}_1(x_o).$$

Whence $P_o \varepsilon F(x_o)$, Q. E. D.

Remark: Note that in proving that (i)$\Rightarrow$(ii) in Theorem 2.1.3 the hypothesis that (X,τ_1) be compact and regular was not required.

THEOREM 2.1.4. *Let* (X,τ) *be a topological space,* $(E,\mathcal{J})$ *be a uniform space, and let* F *be a mapping* $F:X\to\mathcal{C}(E)$. *Then* F *usci at* $a\varepsilon X$ *implies* F *is usc at* $a\varepsilon X$.

Proof: Let $\mathcal{V}(a)$ denote the τ-neighborhood filter at $a\varepsilon X$. Then we must prove

$$\bigcap_{G\varepsilon\mathcal{V}(a)} \overline{F(G)}^{\tau_{\mathcal{J}}} \subset F(a).$$

Suppose it is not true. Then there is a $y\varepsilon\bigcap_{G\varepsilon\mathcal{V}(a)}\overline{F(G)}^{\tau_{\mathcal{J}}}$, such that $y\notin F(a)$. Since $F(a)$ is closed there is a symmetric $J_1\varepsilon\mathcal{J}$, such that $J_1[y]\cap F(a) = \emptyset$. There is a symmetric $J_2\varepsilon\mathcal{J}$ such that $J_2 \circ J_2\subset J_1$. Because F is usci at $a\varepsilon X$ there is a $G_{J_2}\varepsilon\mathcal{V}(a)$ such that $F(G_{J_2})\subset J_2[F(a)]$. Now $y\varepsilon\bigcap_{G\varepsilon\mathcal{V}(a)}\overline{F(G)}^{\tau_{\mathcal{J}}}$ implies $y\varepsilon\overline{F(G_{J_2})}^{\tau_{\mathcal{J}}}$, and thus $J_2[y]\cap F(G_{J_2}) \neq \emptyset$. Thus pick $y_0\varepsilon J_2[y]\cap F(G_{J_2})$, then $(y,y_0)\varepsilon J_2$, $y_0\varepsilon F(x_0)$ for some $x_0\varepsilon G_{J_2}$. Consequently $y_0\varepsilon J_2[F(a)]$ so that there is a $b\varepsilon F(a)$ such that (y_0,b), $(y,y_0)\varepsilon J_2$. Therefore $(y,b)\varepsilon J_2 o J_2\subset J_1$, $b\varepsilon F(a)$, which gives $b\varepsilon J_1[y]\cap F(a)$, a contradiction.

THEOREM 2.1.5. *Let* (X,τ) *be a topological space, let* $(E,\mathcal{J})$ *be a uniform space, let* F *be the mapping* $F:X\to\mathcal{C}(E)$, *and let* $F(X)$ *be relatively compact in* $(E,\mathcal{J})$. *Then* F *is usc at* $a\varepsilon X$ *iff* F *is usci at* $a\varepsilon X$.

Proof: It has already been shown (Theorem 2.1.4) that usci$\Rightarrow$usc. So let F be usc at $a\varepsilon X$, and suppose F is not usci at $a\varepsilon X$. Then

$$\bigcap_{G\varepsilon\mathcal{V}(a)} \overline{F(G)}^{\tau_{\mathcal{J}}} \subset F(a)$$

and there is a $J_1 \varepsilon \mathcal{J}$ such that for every $G\varepsilon\mathcal{V}(a)$ (the τ-neighborhood filter at $a\varepsilon X$) there corresponds an $x_G \varepsilon G$ such that

$$J_1[F(a)] \not\supset F(x_G).$$

Then $\{x_G, G\varepsilon\mathcal{V}(a)\}$ is a net in X and $x_G \xrightarrow[\tau]{} a$. Whence for every $G\varepsilon\mathcal{V}(a)$ there is a $y_G \varepsilon F(x_G)$ such that $y_G \not\varepsilon J_1[F(a)]$. Now $\{y_G, G\varepsilon\mathcal{V}(a)\}$ is a net in $F(X)$. But that $F(X)$ is relatively compact implies there is a subnet, $\{y_{G_\alpha}, \alpha\varepsilon A\}$, of $\{y_G, G\varepsilon\mathcal{V}(a)\}$ such that $y_{G_\alpha} \xrightarrow[\tau_{\mathcal{J}}]{}$ some $y_o \varepsilon E$. We select the corresponding subnet $\{x_{G_\alpha}, \alpha\varepsilon A\}$ of $\{x_G, G\varepsilon\mathcal{V}(a)\}$ and we still have $x_{G_\alpha} \xrightarrow[\tau]{} a$. Whence by Theorem 2.1.3 (X compact, regular was <u>not</u> used in proving 2.1.3 (i)$\Rightarrow$2.1.3 (ii)) $y_o \varepsilon F(a)$. Now there is no loss in generality if we assume J_1 is open in the product topology $\tau_{\mathcal{J}} \times \tau_{\mathcal{J}}$ on $E\times E$. Now J_1 open implies $J_1[F(a)]$ is open. Consequently $E - J_1[F(a)]$ is closed, and we have $y_{G_\alpha} \varepsilon E - J_1[F(a)] =$ closed set, $\alpha\varepsilon A$, $y_{G_\alpha} \xrightarrow[\tau_{\mathcal{J}}]{} y_o$ implies $y_o \not\varepsilon J_1[F(a)]$. But $y_o \varepsilon F(a)$, and we have a contradiction.

THEOREM 2.1.6. Let (X,τ) be a topological space, let $(E,\mathcal{J})$ be a uniform space, let F be a mapping, $F:X\to\mathcal{C}(E)$, and let $F(x)$ be relatively compact in $(E,\mathcal{J})$. If F is lsc at $a\epsilon X$, then F is lsci at $a\epsilon X$.

Proof: Let $\mathcal{V}(a)$ denote the τ-neighborhood filter at $a\epsilon X$. Suppose

$$\overline{\bigcup_{G\epsilon\mathcal{V}(a)} \bigcap_{x\epsilon G} F(x)}^{\tau_{\mathcal{J}}} \supset F(a)$$

and F is not lsci at $a\epsilon X$. Then there is an open $J_1\epsilon\mathcal{J}$ such that for each $G\epsilon\mathcal{V}(a)$

$$J_1[F(x)] \supset F(a), \ x\epsilon G$$

is not true. Whence for each $G\epsilon\mathcal{V}(a)$ there is an $x_G\epsilon G$ and $y_G\epsilon F(a)$ such that $y_G\epsilon E - J_1[F(x_G)]$, which is a closed set. Now $\{y_G, G\epsilon\mathcal{V}(a)\}$ is a net in $F(a)$, and so it has a convergent subnet $\{y_{G_\alpha}, \alpha\epsilon A\}\xrightarrow[\tau_{\mathcal{J}}]{}$ some $y_o\epsilon F(a)$, $\{x_{G_\alpha}, \alpha\epsilon A\}\xrightarrow[\tau]{} a$. Thus let $J\epsilon\mathcal{J}$, and pick $J^*\epsilon\mathcal{J}$ such that $J^*oJ^*\subset J$, J^* symmetric. Since $y_o\epsilon F(a)$, we must have $y_o\epsilon\overline{\bigcup_{G\epsilon\mathcal{V}(a)}\bigcap_{x\epsilon G}F(x)}^{\tau_{\mathcal{J}}}$. Thus $J^*[y_o]\cap\{\bigcup_{G\epsilon\mathcal{V}(a)} \bigcap_{x\epsilon G}F(x)\} \neq \emptyset$. Also since $y_{G_\alpha}\xrightarrow[\tau_{\mathcal{J}}]{} y_o\epsilon F(a)$, there is an $\alpha_o\epsilon A$ such that $\alpha \geq \alpha_o$ implies $y_{G_\alpha}\epsilon J^*[y_o]$. Now there is a $G_o\epsilon\mathcal{V}(a)$ such that $J^*[y_o]\cap\{\bigcap_{x\epsilon G_o} F(x)\} \neq \emptyset$. Let $G_1 = G_{\alpha_o}\cap G_o\subset G_o$, then

$$J^*[y_o]\cap\{\bigcap_{x\epsilon G_1} F(x)\}\supset J^*[y_o]\cap\{\bigcap_{x\epsilon G_o} F(x)\} \neq \emptyset.$$

There is an $\alpha_1 \varepsilon A$ such that $\alpha \geqslant \alpha_1 \geqslant \alpha_o$ implies $x_{G_\alpha} \varepsilon G_1$ and $J^*[y_o] \cap F(x_{G_{\alpha_1}}) \neq \emptyset$. There is a $p \varepsilon F(x_{G_{\alpha_1}})$ such that (p, y_o), $(y_o, y_{G_{\alpha_1}}) \varepsilon J^*$. Thus

$$(p, y_{G_{\alpha_1}}) \varepsilon J^* \circ J^* \subset J, \ p \varepsilon F(x_{G_{\alpha_1}}),$$

which implies that $y_{G_{\alpha_1}} \varepsilon J[p] \subset J[F(x_{G_{\alpha_1}})]$, a contradiction.

The following examples show that we do need to distinguish between usci (lsci) and usc (lsc).

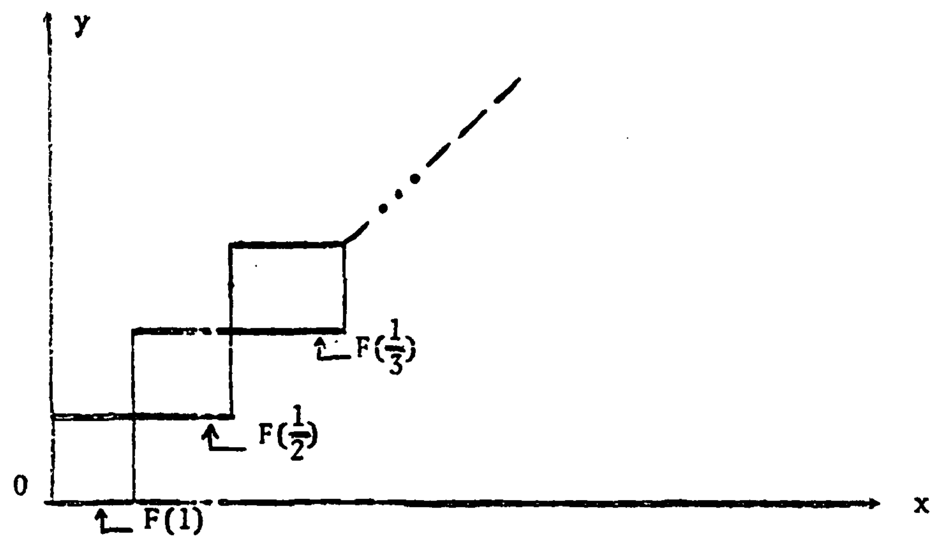
EXAMPLE 2.1.1. Let a mapping $F: [0,1] \to \mathcal{C}(R^2)$ be defined in the following way

$$F(t) = \begin{cases} \{(0,0)\}, & t = 0 \\ SQ(n), & t = 1/n, \ n = 1, 2, 3, \ldots \\ \{(n-1, n-1)\}, & \frac{1}{n} < t < \frac{1}{n-1}, \ n = 2, 3, 4, \ldots \end{cases}$$

where

SQ(n) = the boundary of the unit square having vertices $(n-1, n-1)$, $(n, n-1)$, (n,n), $(n-1,n)$, $n = 1, 2, 3, \ldots$.

We may illustrate the situation graphically.



The location of the other sets $F(t)$ in the xy-plane which correspond to values of t other than those of the form $t = \frac{1}{n}$, $n = 1, 2, 3, \ldots$ is clear. We have that $F(t)$ is a compact subset of R^2 for $0 \leq t \leq 1$. Moreover if we define

$$S_\varepsilon(t_0) = \{t \mid t \in [0,1], \ |t-t_0| < \varepsilon\}, \ \varepsilon > 0,$$

we clearly have

$$\bigcap_{\varepsilon>0} \overline{F(S_\varepsilon(0))} = F(0) = \{(0,0)\} \ ,$$

as a matter of fact

$$\bigcap_{\varepsilon>0} F(S_\varepsilon(t)) = F(t), \forall t \in [0,1].$$

Thus F is usc on $[0,1]$. It is easy to demonstrate that F is not usci at $t = 0$.

EXAMPLE 2.1.2. Let a mapping $F: [0,1] \to \mathcal{C}(R^2)$ be defined as follows

$$F(t) = \begin{cases} \{(x,y) \mid y = 0,\ x \geq 0\} \cup \{(x,y) \mid x = 0,\ y \geq 0\}, \\ \text{if } t \neq 1/n,\ n = 1, 2, 3, \ldots \\ \\ S(n) = \underline{\text{the boundary of the } n\text{-square having vertices}} \\ \qquad (0,0),\ (n,0),\ (n,n),\ (0,n),\ \text{if } t = 1/n, \\ \qquad n = 1, 2, 3, \ldots \end{cases}$$

Then F is usc at $t = 0$, but F is not usci at $t = 0$.

Comment: The concept of the lower semicontinuity of set-valued mappings brings us to an interesting contretemps. For example,

$$C(t) = \{(x,y) \mid x,\ y \in R,\ x^2+y^2 = t^2\},\quad 0 \leq t \leq 1$$

is not lsc at any point of the interval $[0,1]$, but C is continuous on $[0,1]$. However, our interest in semicontinuity in this work stems mainly from the implications of Theorem 2.1.3, and therefore the misbehavior of lower semicontinuity is not critical.

Next we shall point out how the Hausdorff metric [VIII, pg. 166 ff., IX, pg. 131, ex. D, and I, pg. 111 ff.] is related to the uniform structure $2^{\mathcal{U}}$, whenever the uniform structure $\mathcal{U}$ on E is induced by a metric. Let (E,d) be a bounded metric space, $\mathcal{C}^*(E)$ the collection of <u>nonempty</u> closed subsets of E (i.e., $\mathcal{C}^*(E) = \mathcal{C}(E) - \{\emptyset\}$). Define

$$J_\varepsilon = \{(x,y) \mid d(x,y) < \varepsilon,\ x,\ y \in E\},\quad \varepsilon > 0;$$

then $\mathcal{J}_d \equiv [\{J_\varepsilon | \varepsilon>0\}]$ is a uniform structure on E and the topology on E induced by $\mathcal{J}_d$ is the same as the topology on E induced by the metric d. We define a metric on $\mathcal{C}^*(E)$ as follows:

$$\rho(A,B) = \inf \; \{\varepsilon>0 | J_\varepsilon[A] \supset B \;\&\; J_\varepsilon[B] \supset A\}$$

where $A,B\varepsilon\,\mathcal{C}^*(E)$. The details of the proof that ρ is a metric on $\mathcal{C}^*(E)$ are straightforward but tedious and will not be given here. Define

$$\mathbb{N}_\varepsilon = \{(A,B) | A,B\varepsilon\,\mathcal{C}^*(E),\; \rho(A,B)<\varepsilon\}$$

where $\varepsilon>0$, then

$$\mathcal{N} = [\{\mathbb{N}_\varepsilon | \; \varepsilon>0\}]$$

is a uniform structure on $\mathcal{C}^*(E)$. The question which naturally arises is: Does $\mathcal{N} = 2^{\mathcal{J}_d}$? The answer is affirmative. For suppose $\varepsilon>0$ is given, then $\mathbb{N}_\varepsilon \supset \mathbb{V}_{J_\varepsilon/2}$. If $(A,B)\varepsilon\,\mathbb{V}_{J_{\varepsilon/2}}$, then $J_{\varepsilon/2}[A] \supset B$ and $J_{\varepsilon/2}[B] \supset A$. Therefore by the definition of ρ we have that $\rho(A,B)\leq\varepsilon/2<\varepsilon$, which implies $(A,B)\varepsilon\,\mathbb{N}_\varepsilon$. Whence $\mathcal{N} \leq 2^{\mathcal{J}_d}$. Conversely let $\mathbb{V}_{J_\varepsilon}\varepsilon\,2^{\mathcal{J}_d}$, $\varepsilon>0$, then $\mathbb{V}_{J_\varepsilon} \supset \mathbb{N}_\varepsilon$. For if $(A,B)\varepsilon\,\mathbb{N}_\varepsilon$, then $\rho(A,B)<\varepsilon$, which implies that there is a δ, $0<\delta<\varepsilon$ such that $J_\delta[A] \supset B$ and $J_\delta[B] \supset A$. From the fact that $J_\varepsilon[A] \supset J_\delta[A]$ and $J_\varepsilon[B] \supset J_\delta[B]$ we obtain that

$(A,B)\varepsilon \mathcal{V}_{J_\varepsilon}$. Consequently $2^{\gamma_d} \leq \mathcal{N}$, thereby proving $\mathcal{N} = 2^{\gamma_d}$.

Note that the metric ρ is only defined whenever (E,d) is a bounded metric space. This is really no restriction. We define a new metric on E [cf. IX, pg. 31, ex. C] by the equation

$$\bar{d}(x,y) = d(x,y)/[1+d(x,y)], \ x, \ y\varepsilon E.$$

Then $(E,\bar{d})$ is a bounded metric space. Moreover the two topologies induced on E by the metrics d and $\bar{d}$ are the same. We would like much more to be true, viz.,

$$\gamma_d = \gamma_{\bar{d}}.$$

If $\bar{\rho}$ defines the Hausdorff metric on $\mathcal{B}^*(E)$ corresponding to the metric $\bar{d}$, which is bounded, then we would be able to say

$$2^{\gamma_d} = 2^{\gamma_{\bar{d}}},$$

and 2^{γ_d} is uniformly metrizable (with Hausdorff metric $\bar{\rho}$). We state that in fact $\gamma_d = \gamma_{\bar{d}}$ and omit the proof. This motivates the following definition:

DEFINITION 2.1.7. Let (E,d) be an arbitrary metric space, let $\mathcal{B}^*(E)$ denote the nonempty closed subsets of E, and define $\bar{d} \equiv d/(1+d)$,

$$J_\varepsilon = \{(x,y) \mid \bar{d}(x,y) < \varepsilon,\ x,\ y \in E\},\ \varepsilon > 0\ ;$$

then the Hausdorff metric on $\mathcal{C}^*(E)$ is defined by

$$\bar{\rho}(A,B) = \inf \{\varepsilon > 0 \mid J_\varepsilon[A] \supset B\ \&\ J_\varepsilon[B] \supset A\}$$

for A, $B \in \mathcal{C}^*(E)$.

2.2 Convergence of Measurable Functions and Compactness.

Any statement concerning the measurability of sets or of functions is to be understood in terms of ordinary Lebesgue measure, and all integrals are ordinary Lebesgue integrals.

DEFINITION 2.2.1. A function $f: R^p \to R^q$ is measurable iff each component of f is measurable.

In order to eliminate any confusion which might arise regarding two different types of almost everywhere convergence which are common in functional analysis, we state the following definitions.

DEFINITION 2.2.2. Let E be a subset of R^p, let f_n be a sequence of mappings, $f_n: E \to R^q$ $n = 1, 2, 3, \ldots$. Then we say $f_n \to f$ almost everywhere on E as $n \to \infty$ (abbreviated $f_n \to f$ a.e. on E as $n \to \infty$) iff $f_n(t) \to f(t)$ as $n \to \infty$ except possibly on a subset of E of measure zero. Similarly given two mappings f, $g: E \to R^q$ we say $f = g$ almost everywhere on E (abbreviated $f = g$ a.e. on E) iff $f(t) = g(t)$ except possibly on a subset of E of measure zero.

DEFINITION 2.2.3. If $u = (u^1, \ldots, u^q)$, $v = (v^1, \ldots, v^q) \varepsilon R^q$, we define

$$<u,v> = \sum_{i=1}^{q} u^i v^i.$$

DEFINITION 2.2.4. Define

$$L_2^q([a,b]) = \{u \mid u: [a,b] \rightarrow R^q, \int_a^b <u(t),u(t)>dt \text{ \underline{exists and is finite}}\}.$$

Whenever the interval $[a,b]$ is understood we shall simply write L_2^q.

We understand that when speaking of the equality of elements in $L_2^q([a,b])$ that we are actually referring to almost everywhere equality. Defining addition and scalar (real) multiplication in the natural way, L_2^q becomes a real vector space. We define an inner (scalar) product on L_2^q as follows,

(2.2.1) $$(u,v) = \int_a^b <u(t), v(t)>dt,$$

where u, $v \varepsilon L_2^q$. Then $||u|| = \sqrt{(u,u)}$ defines a norm on L_2^q. It is well known that L_2^q with this inner product norm is a real separable Hilbert space. A sequence $u_n \varepsilon L_2^q$, $n = 1, 2, 3, \ldots$ converges to $u \varepsilon L_2^q$ iff

(2.2.2) $$||u_n - u|| \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty,$$

in which case we write

(2.2.3) $$u_n \rightarrow u(st) \quad \text{as} \quad n \rightarrow \infty$$

(to be read: u_n converges strongly to u as $n \rightarrow \infty$).

DEFINITION 2.2.5. A sequence $\{u_n\}$ in $L_2^q([a,b])$ <u>converges weakly to</u> $u \varepsilon L_2^q([a,b])$ iff

(2.2.4) $$(u_n,v) \to (u,v) \text{ as } n \to \infty, \text{ for each } v \varepsilon L_2^q .$$

We shall adopt in this case the notation

(2.2.5) $$u_n \to u \text{ (wk) as } n \to \infty .$$

DEFINITION 2.2.6. Let $\Omega:[a,b] \to \mathbb{P}(R^q)$ = the power set of R^q. Define

$$\mathcal{A}_a^b(\Omega) = \{u \mid u:[a,b] \to R^q,\ u \text{ is measurable on } [a,b],\ u(t) \varepsilon \Omega(t),\ \forall\, t \varepsilon [a,b]\}.$$

THEOREM 2.2.1. (Banach-Saks) <u>Given a sequence</u> u_n <u>in</u> $L_2^m([a,b])$ <u>which converges weakly to</u> $u_o \varepsilon L_2^m([a,b])$, <u>there is a subsequence</u> $\{u_{n_k}\}$ <u>such that the sequence of arithmetic means</u>

$$(u_{n_1}+u_{n_2}+\ldots+u_{n_k})/k \to u_o(st) \text{ as } k \to \infty .$$

Proof: [See XVI, pg. 80].

The following result from real analysis is also required.

THEOREM 2.2.2. <u>Let</u> $\{a_n\}$ <u>be a sequence of real numbers such that</u> $a_n \to a$ as $n \to \infty$. <u>Then</u> $\frac{1}{n}\sum_{i=1}^{n} a_i \to a$ <u>as</u> $n \to \infty$.

Proof: The proof is standard and can be found in many advanced calculus books.

THEOREM 2.2.3. Let Ω be a mapping, $\Omega:[a,b]\to\mathcal{C}\#(R^m)$, where $\mathcal{C}\#(R^m)$ denotes the collection of compact (nonempty) subsets of R^m. If $\bigcup_{a\leq t\leq b}\Omega(t)$ is bounded and Ω is usc on $[a,b]$, then $\mathcal{A}_a^b(\Omega)\neq\emptyset$.

Proof: The proof given for this theorem is patterned after the proof of Filippov's lemma in [7, pp. 78-79], although this theorem is not a consequence of Filippov's lemma. In order to prove the theorem a measurable function $u:[a,b]\to R^m$ must be constructed such that $u(t)\in\Omega(t)$, $a\leq t\leq b$. Since $\Omega(t)$ is compact, there is a $u\in\Omega(t)$ such that its first coordinate u^1 = minimum on $\Omega(t)$. If there is more than one such, choose u so that its first two coordinates u^1, u^2=minimum on $\Omega(t)$, etc. In this way a function $u:[a,b]\to R^m$ is defined such that $u(t)\in\Omega(t)$, $a\leq t\leq b$. The proof of the theorem is concluded by demonstrating that u is measurable on $[a,b]$. The method of proof is by mathematical induction. Let $c\in R$ and define

$$F_c = \{t\,|\,a\leq t\leq b,\ u^1(t)\leq c\}.$$

Then F_c is closed. For suppose this is not the case. Then there is a sequence $\{t_n\}$ in F_c such that $t_n\to t_o$ as $n\to\infty$, but $t_o\notin F_c$, i.e., $t_n\in F_c$, $n = 1, 2, 3, \ldots$, $t_n\to t_o$, and $u^1(t_o)>c$. Consequently there is an $\varepsilon_1>0$ such that $u^1(t_n)\leq u^1(t_o)-\varepsilon_1$ for $n = 1, 2, 3, \ldots$. Now $\{u(t_n)\}$ is a bounded sequence in R^m. Thus there is a subsequence of $\{u(t_n)\}$ (still denoted by $\{u(t_n)\}$) such that $u(t_n)\to\tilde{u}$ and $t_n\to t_o$ as $n\to\infty$. It follows from Theorem 2.1.3 that $\tilde{u}\in\Omega(t_o)$.

Moreover

$$\lim u^1(t_n) = \tilde{u}^1 \leq u^1(t_o) - \varepsilon_1,$$

which contradicts the definition of $u^1(t_o)$. Thus $c\varepsilon R$ implies F_c is closed, and thus u^1 is measurable on $[a,b]$. Assume $u^1, \ldots, u^k$ are measurable on $[a,b]$ for $1 \leq k < m$. Then by Lusin's theorem [XIII, pg. 236] it follows that for any $\varepsilon > 0$ there is a closed set $E_\varepsilon \subset [a,b]$ such that $u^1, \ldots, u^k$ are continuous on E_ε and the measure of E_ε is greater than $(b-a)-\varepsilon$. The set $\{t\varepsilon E_\varepsilon | u^{k+1}(t) \leq c\}$, $c\varepsilon R$ is closed. If this were not the case, then there would be a sequence $t_n \varepsilon E_\varepsilon$, $u^{k+1}(t_n) \leq c$, $n = 1, 2, 3, \ldots$, $t_n \to t_o$ as $n \to \infty$ and $u^{k+1}(t_o) > c$. Therefore there is an $\varepsilon_1 > 0$ such that $u^{k+1}(t_n) \leq u^{k+1}(t_o) - \varepsilon_1$ $n = 1, 2, 3, \ldots$, $t_n \to t_o \varepsilon E_\varepsilon$ as $n \to \infty$. Since $\{u(t_n)\}$ is bounded there is a subsequence of $\{u(t_n)\}$ (still called $\{u(t_n)\}$) such that $u(t_n) \to \tilde{u}$, $t_n \to t_o$ as $n \to \infty$. Consequently by Theorem 2.1.3 $\tilde{u} = (\tilde{u}^1, \ldots, \tilde{u}^m)$ $\varepsilon\Omega(t_o)$. Now using the continuity of the u^i, $i = 1, \ldots, k$ on E_ε we have $\tilde{u}^i = u^i(t_o)$, $i = 1, 2, \ldots, k$, whereas $\tilde{u}^{k+1} \leq u^{k+1}(t_o) - \varepsilon_1$. This contradicts the definition of $u^{k+1}(t_o)$. Therefore F_c is closed, $c\varepsilon R$, and u^{k+1} is measurable on E_ε (which has measure greater than $b-a-\varepsilon$) and $\varepsilon > 0$ is arbitrary. It follows that u^{k+1} is measurable on the whole interval $[a,b]$. Whence $u = (u^1, \ldots, u^m)$ is measurable on $[a,b]$ and $\mathcal{A}_a^b(\Omega) \neq \emptyset$.

THEOREM 2.2.4. Let Ω be a mapping $\Omega:[a,b]\to \mathcal{C}\#(R^m)$ (defined in Theorem 2.2.3). If $\bigcup_{a\leq t\leq b}\Omega(t)$ is bounded and $\Omega(t)$ is convex, $a\leq t\leq b$, then $\mathcal{A}_a^b(\Omega)$ is weakly compact in itself (i.e., any sequence $\{u_n\}$ in $\mathcal{A}_a^b(\Omega)$ has a subsequence $\{u_{n_k}\}$ such that $u_{n_k}\to u(wk)$ as $k\to\infty$ for some $u\varepsilon\mathcal{A}_a^b(\Omega)$).

Proof: Each u in $\mathcal{A}_a^b(\Omega)$ is bounded and measurable on $[a,b]$, and thus belongs to $L_2^m([a,b])$. Whence $\mathcal{A}_a^b(\Omega)$ is a bounded subset of $L_2^m([a,b])$. Let $u_n\varepsilon\mathcal{A}_a^b(\ )$, $n = 1, 2, 3, \ldots,$ then by a well known theorem [XII, pg. 117] there is a subsequence of $\{u_n\}$ (still denoted by $\{u_n\}$) such that

$$u_n\to u(wk) \quad \text{as} \quad n\to\infty \quad \text{for some} \quad u\varepsilon L_2^m([a,b]).$$

What we must prove is that $u\varepsilon\mathcal{A}_a^b(\ )$. By Theorem 2.2.1 there is a further subsequence of $\{u_n\}$ (still called $\{u_n\}$) such that $\sigma_n = \frac{1}{n}\sum_{i=1}^{n} u_i\to u(st)$ as $n\to\infty$. Therefore (see [XI, pg. 87]) there is a subsequence of $\{\sigma_n\}$ (say $\{\sigma_{n_k}\}$) such that

$$\sigma_{n_k}\to u \text{ (a.e.) on } [a,b] \text{ as } n\to\infty. \tag{2.2.6}$$

Note that $\sigma_{n_k}(t)$ is just a convex linear combination of points in $\Omega(t)$, $a\leq t\leq b$. Since $\Omega(t)$ is convex for $a\leq t\leq b$, it follows that $\sigma_{n_k}(t)$ is just a sequence of points in $\Omega(t)$, $a\leq t\leq b$. Thus by

(2.2.6) $u(t) \in \Omega(t)$ for almost every $t \in [a,b]$. On the possible exceptional set of measure zero at which $u(t) \notin \Omega(t)$, we can redefine u such that $u(t) \in \Omega(t)$ for each t on the interval $[a,b]$. Thus $u_n \to u(wk)$ as $n \to \infty$, $u \in \mathcal{A}_a^b(\Omega)$ (recall that functions are equal in L_2^m, if they are equal a.e.).

The special case of this theorem where $\Omega(t) = \Omega(a)$, for each $t \in [a,b]$ was proved by Lee and Markus [11, pg. 39]. The method used here, however, differs substantially from that of Lee and Markus.

CHAPTER III

THE EXISTENCE OF OPTIMAL CONTROLS FOR LINEAR SYSTEMS

3.1 Formulation of the Linear Control Problem

We confine attention in this chapter to an optimal control problem in which the control system at any time t can be described by a system of real ordinary differential equations of the following type,

$$\dot{x}(t) = A(t)x(t)+B(t)u(t)+\omega(t), \tag{3.1.1}$$

where for each $t\varepsilon[0,T]$, $T>0$ fixed,

$x(t)$ is an nx1 real matrix;

$A(t)$ is an nxn real matrix;

$B(t)$ is an nxm real matrix;

$u(t)$ is an mx1 real matrix;

$\omega(t)$ is an nx1 real matrix.

We shall not use any notational device to distinguish between row and column vectors. The context should make it clear which is meant.

The symbol u appearing in (3.1.1) denotes a bounded measurable control function with domain contained in $[0,T]$, and with range R^m. Suppose $\Omega:[0,T]\rightarrow\mathcal{C}(R^m)$, and $\bigcup_{0\leq t\leq T}\Omega(t)$ is bounded.

Let Γ be a closed subset of $[0,T]$. We define a set $\mathcal{U}(\Omega,\Gamma)$ by the relation

$$\mathcal{U}(\Omega,\Gamma) = \bigcup_{t\in\Gamma} \mathcal{A}_0^t(\Omega)\times\{t\} \tag{3.1.2}$$

(see Definition 2.2.6 for the meaning of $\mathcal{A}_0^t(\Omega)$). If the matrices A, B, ω, appearing in (3.1.1), are continuous on $[0,T]$, then given $x_0\varepsilon R^n$, $(u,t_1)\varepsilon\mathcal{U}(\Omega,\Gamma)$ there is a unique absolutely continuous function $x(\cdot,u):[0,t_1]\to R^n$ satisfying (3.1.1) a.e. on $[0,t_1]$, and the initial condition

$$x(0,u) = x_0, \tag{3.1.3}$$

(see [VI, pg. 74 ff.]). In this case $x(\cdot,u):[0,t_1]\to R^n$ will be called the response to the control u. By the method of variation of parameters [VI, pg. 74ff.] the response to the control $(u,t_1)\varepsilon\mathcal{U}(\Omega,\Gamma)$ is that function defined by the relation

$$x(t,u) = X(t)[x_0+\int_0^t X^{-1}(\xi)[B(\xi)u(\xi)+\omega(\xi)]d\xi] \tag{3.1.4}$$

for $0\leq t\leq t_1$, where the nxn matrix function X is defined by the matrix differential equation

$$\dot{X}(t) = A(t)X(t), \quad t\geq 0; \quad X(0) = I_n,$$

and I_n is the nxn identity matrix.

Let the target function $F:\Gamma \to \mathcal{C}(R^n)$ be such that $F(t) \neq \emptyset$ for every $t \in \Gamma$. We require of F that it be usc on Γ. We now define $\hat{\mathcal{U}}(\Omega,\Gamma)$ to be that particular subset of $\mathcal{U}(\Omega,\Gamma)$ defined by the following condition

(3.1.5) $\qquad (u,t_1) \in \hat{\mathcal{U}}(\Omega,\Gamma)$ iff $(u,t_1) \in \mathcal{A}_0^{t_1}(\Omega) \times \{t_1\}$

and $x(t_1,u) \in F(t_1)$.

If $(u,t_1) \in \hat{\mathcal{U}}(\Omega,\Gamma)$, we shall say (u,t_1) is an <u>admissible control function</u>, and in the event that the domain, $[0,t_1]$, $t_1 \in \Gamma$, of u is specified we shall also say that $u: [0,t_1] \to R^m$ is an admissible control function (admissible steering function, or simply an admissible control).

Finally suppose

(3.1.6) $\mathcal{k}:[0,T]\times R^n \times R^m \to R$ is lower semicontinuous (lsc), and satisfies $|\mathcal{k}(t,x,u)| \leq \mu(t) g(||x||)$, where μ is summable on any finite interval, and $g(||x||) = O(||x||)$ as $||x|| \to \infty$.

Utilizing the mapping in (3.1.6), we define the functional $K:\mathcal{U}(\Omega,\Gamma) \to R$ by the equation

(3.1.7) $$K(u,t_1) = \int_0^{t_1} \mathcal{k}(\xi,\ x(\xi,u),\ u(\xi))\,d\xi,$$

when $(u,t_1) \in \mathcal{U}(\Omega,\Gamma)$.

The existence of the integral in (3.1.7) can be proved by using a theorem in McShane [XIII, pg. 123]. The mapping K in (3.1.7) is called the <u>cost functional</u> (the criterion of optimality or the objective) in the literature. The optimal control problem to be considered in this chapter is

$$(3.1.8) \qquad K(u,t_1) = \text{minimum on } \hat{\mathcal{U}}(\Omega,\Gamma).$$

Certain special cases of this problem have been investigated with varying degrees of thoroughness by a number of authors [1, 3, 10, 11, 13, and 17], but a study of problem (3.1.8) in the generality of its present formulation does not appear to have been attempted. Lee and Markus [11] do, however, show that if

$$\mathcal{k}(t,x,u) = a(t,x)+b_i(t,x)u^i$$

(with summation on $i = 1, 2, \ldots, m$), a, b_i are continuous $i = 1, 2, \ldots, m$, Ω is a fixed compact convex subset of R^m, $\Gamma = [0,T]$, $F(t)$ compact for $t\in[0,T]$, $F:[0,T]\to\mathcal{C}(R^n)$ is continuous, and $\hat{\mathcal{U}}(\Omega,\Gamma)\neq\emptyset$, then a solution to (3.1.8) exists. In [10] the time optimal control problem ($\mathcal{k}\equiv 1$, Ω the unit cube in R^m) for a linear control system was studied in great detail. Balakrishnan in [1] studied the problem of minimizing $||x(t,u)-x_1||$ on $\mathcal{B}$, a closed and bounded convex subset of $L_2^m([0,T])$, x_1 a fixed point in R^n. Roxin's existence theorem [17] covers (3.1.8), if for each fixed $(t,x)\in[0,T]\times R^n$, $H(t,x,\Omega)$ is a <u>compact convex subset</u> of R^{n+1}, Ω is a fixed compact subset of R^m, and we define

$$H(t,x,u) = (A(t)x(t)+B(t)u(t)+\omega(t), \mathscr{k}(t,x,u)),$$

i.e., $H(t,x,u)\in R^{n+1}$, and the first n components of $H(t,x,u)$ are given by $A(t)x(t)+B(t)u(t)+\omega(t)$, while the $(n+1)\underline{st}$ component of $H(t,x,u)$ is $\mathscr{k}(t,x,u)$ (actually Roxin in [20] also assumes $\mathscr{k}$ is continuous in (x,u) for each fixed t, $\mathscr{k}$ is summable with respect to t for each fixed (x,u), and $\mathscr{k}$ is Lipschitzian with respect to x and satisfies (3.1.6)). The excellent contribution by Neustadt [13] has removed all assumptions of convexity from both the system (3.1.1) and the restraint set Ω. But Neustadt's existence theorem covers (3.1.8) only if very special restrictions are placed on the mapping $\mathscr{k}$ in (3.1.6), viz.,

$$\mathscr{k}(t,x,u) = \langle\alpha(t),x(t)\rangle+\phi^o(u,t)$$

where $\alpha:[0,T]\to R^n$, $\phi^o:R^m\times[0,T]\to R$ are both continuous. Finally there is the recent paper of Chang [3]. Chang assumes that the matrices A, B, and ω in (3.1.1) are constant (actually $\omega = 0$). He assumes also that the cost functional in (3.1.7) has the form

$$K(u) = \int_0^\infty \langle x(t,u), Qx(t,u)\rangle+\langle u(t), Ru(t)\rangle dt$$

where Q is a non-negative definite $m\times m$ matrix of real constants, and T is an $m\times m$ positive definite matrix of real constants. The restraint set Ω is the "unit cube" in R^m (the domain of each admissible control function is $[0,\infty)$). Under these hypotheses Chang is able to prove there exists an admissible control for which the function K achieves a minimum (the class of admissible controls is the collection of bounded measurable functions having their range in the "unit cube" of R^m, and which transfer an initial point $x_1 \varepsilon R^n$ to the origin as $t \to \infty$). This result is related to some of the theorems in this chapter. Especially Chang has anticipated our use of the Banach-Saks theorem (Theorem 2.2.1). The results in this chapter, however, were obtained (and presented in a seminar on variational theory under the dual sponsorship of Dr. G. M. Ewing and Dr. W. T. Reid in September, 1964) before the results of Chang were published or otherwise known to the author.

The present hypotheses for problem (3.1.8) are not restrictive enough to yield an existence theorem for the problem. The subject of this chapter then is to choose judiciously a few more hypotheses which will enable us to prove that a solution to (3.1.8) does exist.

We now prove the fundamental compactness theorem for this chapter, but first we need the following definition.

DEFINITION 3.1.1 $\hat{\mathcal{U}}(\Omega,\Gamma)$ (see (3.1.5) for terminology) is <u>weakly compact in itself</u> iff for any sequence $(u_n,t_n) \varepsilon \hat{\mathcal{U}}(\Omega,\Gamma)$, $n = 1, 2, 3, \ldots$, there is a $t_o \varepsilon \Gamma$ and a $u_o \varepsilon L_2^m([0,t_o])$ such that some subsequence $\{(u_{n_j},t_{n_j})\}$ of $\{(u_n,t_n)\}$ has the property

$$\tilde{u}_{n_j} \to u_o (wk),\ t_{n_j} \to t_o \quad \text{as} \quad j \to \infty ,$$

and $(u_o,t_o) \varepsilon \hat{\mathcal{U}}(\Omega,\Gamma)$, where $\tilde{u}_{n_j} \varepsilon\ L_2^m([0,t_o])$ is defined by

$$\tilde{u}_{n_j} = \begin{cases} u_{n_j}|[0,t_o] & \text{if } t_o \leq t_{n_j} \\ \bar{u}_{n_j} & \text{if } t_{n_j} < t_o \end{cases}$$

$j = 1, 2, 3, \ldots,$ where $u_{n_j}|[0,t_o]$ denotes the restriction of u_{n_j} to $[0,t_o]$, and $\bar{u}_{n_j}$ is given by

$$\bar{u}_{n_j}(t) = \begin{cases} u_{n_j}(t), & \text{if } 0 \leq t \leq t_{n_j} \\ u^*(t), & \text{if } t_{n_j} < t \leq t_o, \end{cases}$$

u^* any function in $\mathcal{A}_0^{t_o}(\Omega)$.

THEOREM 3.1.1. *Let* Ω *be a mapping,* $\Omega: [0,T] \to \mathcal{B}^{\#}(R^m)$ *(see Theorem 2.2.3 for the meaning of* $\mathcal{B}^{\#}(R^m)$*). If* $\bigcup_{0 \leq t \leq T} \Omega(t)$ *is bounded, if* Ω *is usc on* $[0,T]$, Γ *is closed, and if* $\Omega(t)$ *is convex for* $0 \leq t \leq T$, *then* $\hat{\mathcal{U}}(\Omega,\Gamma)$ *is weakly compact in itself.*

Proof: Let $(u_n,t_n) \in \hat{\mathcal{U}}(\Omega,\Gamma)$, $n = 1, 2, 3, \ldots,$ then since $t_n \in [0,T]$, $n = 1, 2, 3, \ldots,$ there is a subsequence of $\{(u_n,t_n)\}$ (still called $\{(u_n,t_n)\}$) such that $t_n \to t_o$ as $n \to \infty$. Moreover, the convergence of t_n to t_o can be assumed to be monotone. Now since Γ is a closed subset of $[0,T]$ and $t_n \varepsilon \Gamma$, $n = 1, 2, 3, \ldots$, it follows that $t_o \varepsilon \Gamma$. There are two cases:

Case 1 $\qquad 0 \leq t_o \leq t_n \leq T,\ n = 1, 2, 3, \ldots;$

Case 2 $\qquad 0 \leq t_n \leq t_o \leq T,\ n = 1, 2, 3, \ldots.$

We consider Case 1 first. Applying Theorem 2.2.4 we are able to conclude there is a further subsequence of $\{(u_n, t_n)\}$ (still called $\{(u_n, t_n)\}$) such that $u_n \to u_o(wk)$ as $n \to \infty$ for some $u_o \varepsilon \mathcal{A}_0^{t_o}(\Omega)$. More precisely denoting the restriction of u_n to $[0,t_o]$ by $u_n|[0,t_o]$, we have

$$u_n|[0,t_o] \to u_o(wk),\ t_n \to t_o \text{ as } n \to \infty,\ t_o \varepsilon \Gamma,\ u_o \varepsilon \mathcal{A}_0^{t_o}(\Omega). \tag{3.1.9}$$

It remains to show that the response to (u_o, t_o), viz., $x(\cdot, u_o)$ satisfies

$$x(t_o, u_o) \varepsilon F(t_o). \tag{3.1.10}$$

(cf. The definition of $\hat{\mathcal{U}}(\Omega, \Gamma)$ in (3.1.5)). Consulting equation (3.1.4) we find that

$$x(t, u_n) = X(t)[x_o + \int_0^t X^{-1}(\xi)[B(\xi)u_n(\xi) + \omega(\xi)]d\xi],\ 0 \leq t \leq t_n. \tag{3.1.11}$$

It follows from (3.1.9) and (3.1.11) that

$$x(t, u_n) \to x(t, u_o) \text{ as } n \to \infty,\ 0 \leq t \leq t_o. \tag{3.1.12}$$

Utilizing (3.1.11) we obtain the inequality

$$||x(t_n,u_n)-x(t_o,u_n)||\leq||X(t_n)-X(t_o)|| \; +$$

$$||X(t_n)-X(t_o)||\int_0^{t_o}||X^{-1}(\xi)[B(\xi)u_n(\xi)+\omega(\xi)]||d\xi \; +$$

$$||X(t_n)-X(t_o)||\int_{t_o}^{t_n}||X^{-1}(\xi)[B(\xi)u_n(\xi)+\omega(\xi)]||d\xi \; .$$

All admissible controls are bounded (in R^m) and B, X, X^{-1}, ω are all continuous on $[0,T]$, and thus the integrands appearing on the right hand side of the last inequality are bounded. Since $t_n \to t_o$ as $n \to \infty$ the right hand side of the last inequality is a null sequence. Whence

(3.2.13) $$[x(t_n,u_n)-x(t_o,u_n)] \to 0 \quad \text{as} \quad n \to \infty \; .$$

Consequently from (3.1.13) and (3.1.12) we obtain that

(3.1.14) $$x(t_o,u_o) = \lim[x(t_n,u_n)-x(t_n,u_n)+x(t_o,u_n)]$$

$$= \lim x(t_n,u_n) = \lim x(t_o,u_n) = x(t_o,u_o).$$

Now since $(u_n,t_n)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$, $n = 1, 2, 3, \ldots,$ it follows from the definition of $\hat{\mathcal{U}}(\Omega,\Gamma)$ that

(3.1.15) $$x(t_n,u_n)\varepsilon F(t_n), \; n = 1, 2, 3, \ldots \; .$$

Whence from (3.1.14), (3.1.15), the fact that $F:\Gamma \to \mathcal{C}(R^n)$ is usc, and $t_n \to t_o$ as $n \to \infty$, we have (by Theorem 2.1.3) that $x(t_o,u_o)\varepsilon F(t_o)$. Therefore $(u_o,t_o)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$ and this completes the proof for Case 1.

In the event Case 2 obtains, we extend each control in the original sequence $\{(u_n, t_n)\}$ to the entire interval $[0,t_o] \supset [0,t_n]$ in an appropriate manner. We have by Theorem 2.2.3 that $\mathcal{A}_0^{t_o}(\Omega) \neq \emptyset$. Select $u^* \varepsilon \mathcal{A}_0^{t_o}(\Omega)$, and define

$$\bar{u}_n(t) = \begin{cases} u_n(t), & \text{if } 0 \leq t \leq t_n \\ u^*(t), & \text{if } t_n < t \leq t_o \end{cases} .$$

Then by replacing u_n by $\bar{u}_n$ in the argument made for Case 1, we find that a subsequence of $\{(\bar{u}_n, t_n)\}$ (still called $\{(\bar{u}_n, t_n)\}$) has the property that

$$\bar{u}_n \rightarrow u_o(wk),\ t_n \rightarrow t_o \quad \text{as} \quad n \rightarrow \infty,\ t_o \varepsilon \Gamma,\ u_o \varepsilon \mathcal{A}_0^{t_o}(\Omega).$$

We can also prove in a manner quite similar to that used in Case 1 that $x(t_o, u_o) \varepsilon F(t_o)$, and thus $(u_o, t_o) \varepsilon \hat{\mathcal{U}}(\Omega,\Gamma)$ which completes the proof.

3.2 Existence Theorems for the Linear Optimal Control Problem within the Class $\hat{\mathcal{U}}(\Omega,\Gamma)$.

THEOREM 3.2.1. Let the hypotheses of Theorem 3.1.1 and Hypothesis (3.1.6) remain in effect. In addition let the following three hypotheses be satisfied:

(i) $\hat{\mathcal{U}}(\Omega,\Gamma) \neq \emptyset$;

(ii) For any $(t,x) \varepsilon [0,T] \times R^n$, the function k introduced in (3.1.6) is convex in the variable $u \varepsilon R^m$, i.e., $p,\ q \geq 0$ $p+q = 1$,

$u, v \in R^m$ *imply*

$$\mathcal{k}(t,x,pu+qv) \leq p\,\mathcal{k}(t,x,u)+q\,\mathcal{k}(t,x,v)$$

(actually the special case $p = q = \frac{1}{2}$ *is all that is required in the proof of this theorem)*;

(iii) $\mathcal{k}$ *satisfies a Lipschitz condition in the variable* x, *i.e.*, *there is a constant* $A>0$ *such that whenever* $(t,u) \in [0,T]\times[\bigcup_{0\leq t\leq T}\Omega(t)]$, $x, y \in R^n$, *we have*

$$||\mathcal{k}(t,x,u)-\mathcal{k}(t,y,u)|| \leq A||x-y||.$$

Then there is a $(u_0,t_0) \in \hat{\mathcal{U}}(\Omega,\Gamma)$ *such that*

$$K(u_0,t_0) = \inf K(\hat{\mathcal{U}}(\Omega,\Gamma)).$$

Proof: We first observe that the set

$$\mathcal{B} = \{x(t,u) \mid (u,t_1) \in \hat{\mathcal{U}}(\Omega,\Gamma),\ 0\leq t\leq t_1\}$$

is a bounded subset of R^n. This follows immediately from eq. (3.1.4), if one notices X, X^{-1}, B, ω are all continuous on $[0,T]$ and that $\bigcup_{0\leq t\leq T}\Omega(t)$ is bounded. Consequently since $\mathcal{k}$ is lsc on $[0,T]\times R^n\times R^m$ it follows that

$$\mathcal{k}([0,T]\times\mathcal{B}\times\bigcup_{0\leq t\leq T}\Omega(t))$$

is bounded from below. Therefore by hypothesis (i) $K(\hat{\mathcal{U}}(\Omega,\Gamma))$

is a nonempty set of real numbers bounded below. Whence we obtain that

$$+\infty > \inf K\,(\hat{\mathcal{U}}(\Omega,\Gamma)) \overset{=}{d} \gamma > -\infty.$$

There is a "minimizing sequence" of controls

$$(3.2.1) \qquad \{(u_n,t_n)\} \subset \hat{\mathcal{U}}(\Omega,\Gamma),\ K(u_n,t_n) \to \gamma \quad \text{as} \quad n \to \infty.$$

Clearly this "minimizing sequence" admits a subsequence (still called $\{(u_n,t_n)\}$) such that $t_n \to$ some $t_o \varepsilon \Gamma$ monotonely as $n \to \infty$.

There are again two cases:

Case 1 $\qquad 0 \leq t_o \leq t_n \leq T,\ n = 1, 2, 3, \ldots,$

Case 2 $\qquad 0 \leq t_n \leq t_o \leq T,\ n = 1, 2, 3, \ldots,$

We exhibit the details of the proof for Case 1 first. In view of Theorem 3.1.1, there is a further subsequence of $\{(u_n,t_n)\}$ (still denoted by $\{(u_n,t_n)\}$) such that

$$(3.2.2) \qquad u_n \to u_o(wk),\ t_n \to t_o \quad \text{as} \quad n \to \infty,\ (u_o,t_o) \varepsilon\ \hat{\mathcal{U}}(\Omega,\Gamma).$$

Whence by the Banach-Saks theorem (Theorem 2.2.1) there is a further subsequence of $\{(u_n,t_n)\}$ (without changing the notation) such that

$$(3.2.3) \qquad \sigma_n = \frac{1}{n}\sum_{i=1}^{n} u_i \to u_o(st),\ t_n \to t_o \quad \text{as} \quad n \to \infty.$$

Thus (see [XI, pg. 87]) there is a subsequence of $\{(\sigma_n,t_n)\}$, say $\{(\sigma_{n_j}, t_{n_j})\}$ such that

(3.2.4) $\sigma_{n_j} \to u_o$ (a.e.) on $[0,t_o]$ as $j\to\infty$.

From equation (3.1.4) and the fact that $u_n \to u_o$ (wk) as $n\to\infty$ we have that

(3.2.5) $x(t,u_n)\to x(t,u)$ as $n\to\infty$, $0\leq t\leq t_o$.

It thereby follows from Theorem 2.2.2 that

(3.2.6) $$\lim_{j\to\infty} \frac{1}{n_j} \sum_{i=1}^{n_j} x(t,u_i) = \lim_{n\to\infty} x(t,u_n) = x(t,u_o),\ 0\leq t\leq t_o \ .$$

Since $\mathcal{k}$ is convex in u (hypothesis (ii) of this theorem) we have that (see eq. (3.2.3))

(3.2.7) $$\mathcal{k}(\xi,x(\xi,u_{n_j}),\sigma_{n_j}(\xi))\leq\frac{1}{n_j} \sum_{i=1}^{n_j} \mathcal{k}(\xi,x(\xi,u_{n_j}),u_i(\xi))$$

for $0\leq\xi\leq t_o$, $j = 1, 2, 3, \ldots$. Also the following equalities are clearly valid:

(3.2.8) (a) $$\lim_{n\to\infty} K(u_n,t_n) = \lim_{j\to\infty} K(u_{n_j},t_{n_j}) = \gamma;$$

(b) $$\lim_{j\to\infty} K(u_{n_j},t_{n_j}) = \lim_{j\to\infty} \int_0^{t_{n_j}} \mathcal{k}(\xi,x(\xi,u_{n_j}),u_{n_j}(\xi))d\xi$$

$$= \lim_{j\to\infty} \int_0^{t_o} \mathcal{k}(\xi,x(\xi,u_{n_j}),u_{n_j}(\xi))d\xi$$

$$= \lim_{n\to\infty} \int_0^{t_o} \mathcal{k}(\xi, x(\xi,u_n),u_n(\xi))d\xi$$

Relation (3.2.8) (a) is evident, whereas (3.2.8) (b) is an easy consequence of hypothesis (3.1.6)

The next step in the proof is to show that

$$(3.2.9)\qquad \frac{1}{n_j}\sum_{i=1}^{n_j}\int_0^{t_o} k(\xi,x(\xi,u_{n_j}),u_i(\xi))d\xi \to \gamma \quad \text{as} \quad j\to\infty.$$

We observe that by Theorem 2.2.2, eq. (3.2.8) (a), (b) we obtain

$$(3.2.10)\qquad \lim_{j\to\infty}\frac{1}{n_j}\sum_{i=1}^{n_j}\int_0^{t_o} k(\xi,x(\xi,u_i),u_i(\xi))d\xi = \gamma.$$

Thus in order to prove (3.2.9) it will suffice to prove

$$(3.2.11)\qquad \lim_{j\to\infty}\frac{1}{n_j}\sum_{i=1}^{n_j}\int_0^{t_o}\{k(\xi,x(\xi,u_{n_j}),u_i(\xi))- k(\xi,x(\xi,u_i),u_i(\xi))\}d\xi = 0.$$

From hypothesis (iii) of this theorem we obtain the inequality

$$(3.2.12)\qquad \left|\frac{1}{n_j}\sum_{i=1}^{n_j}\int_0^{t_o}\{k(\xi,x(\xi,u_{n_j}),u_i(\xi))- k(\xi,x(\xi,u_i)u_i(\xi))\}d\xi\right| \leq$$

$$\frac{1}{n_j}\sum_{i=1}^{n_j} A\int_0^{t_o} ||x(\xi,u_{n_j})-x(\xi,u_i)||d\xi.$$

We have that

$$\lim_{j\to\infty} x(\xi,u_{n_j}) = \lim_{i\to\infty} x(\xi,u_i) = x(\xi,u_o),\ 0\leq\xi\leq t_o,$$

and it follows from the Lebesgue dominated covergence theorem (or even the bounded convergence theorem) that

$$\lim_{i\to\infty} \int_0^{t} ||x(\xi, u_i) - x(\xi,u_o)||d\xi = 0$$

and

$$\lim_{j\to\infty} \int_0^{t_o} ||x(\xi,u_{n_j}) - x(\xi, u_o)||d\xi = 0.$$

Thus given $\varepsilon>0$ there is a positive integer n_ε^o, which can be taken to be one of the integers in the collection $n_1<n_2<n_3<\ldots<n_j<\ldots$, such that

(3.2.13) $$n_j, i \geq n_\varepsilon^o \Rightarrow A \int_0^{t_o} ||x(\xi,u_{n_j})-x(\xi,u_i)||d\xi<\varepsilon/2.$$

This n_ε^o is fixed (it depends only on the given $\varepsilon>0$). Now since $x(\xi,u_{n_j}),x(\xi,u_i)$, $i, j = 1, 2, 3, \ldots$ are uniformly bounded on $[0,t_o]$, it follows that there is a $\lambda>0$ such that

(3.2.14) $$0\leq \sum_{i=1}^{n_\varepsilon^o} \int_0^{t_o} A||x(\xi,u_{n_j})-x(\xi,u_i)||d\xi\leq \lambda ,$$

for all n_j, $j = 1, 2, 3, \ldots$. There is an integer $n_\varepsilon>n_\varepsilon^o$ such that

(3.2.15) $$n_j\geq n_\varepsilon>n_\varepsilon^o \Longrightarrow \lambda/n_j<\varepsilon/2 .$$

Denote the sum on the right hand side of inequality (3.2.12) by S_j. Then given $n_j\geq n_\varepsilon$, it follows that

$$(\leq S_j = \frac{1}{n_j}\sum_{i=1}^{n_\varepsilon^o}\int_0^{t_o} A||x(\xi,u_{n_j})-x(\xi,u_i)||d\xi + \frac{1}{n_j}\sum_{i=n_\varepsilon^o+1}^{n_j}\int_0^{t_o} A||x(\xi,u_{n_j})-x(\xi,u_i)||d\xi \ .$$

Consequently by (3.2.13), (3.2.14), and (3.2.15) one obtains that

$$(3.2.16) \qquad n_j \geq n > n_\varepsilon^o \quad 0 \leq S_j \leq \varepsilon/2 + \frac{1}{n_j}\sum_{i=n_\varepsilon^o+1}^{n_j}\{\varepsilon/2\} < \varepsilon \ .$$

Therefore $S_j \to 0$ as $j \to \infty$, and (3.2.11) is true in view of inequality (3.2.12). This establishes the validity of (3.2.9).

Returning to inequality (3.2.7) we recall that $\bigcup_{0 \leq t \leq T}\Omega(t)$ is bounded, $\sigma_{n_j}(t) \in \Omega(t)$, for $0 \leq t \leq t_o \leq T$, and the responses $x(\cdot,u_n)$, $n = 1, 2, 3, \ldots$ are uniformly bounded on $[0,t_o]$. It follows then from the lower semicontinuity of $\mathcal{k}$ (hypothesis 3.1.6) that there is a real number α such that

$$(3.2.17) \qquad \mathcal{k}(\xi,x(\xi,u_{n_j}),\sigma_{n_j}(\xi)) \geq \alpha, \quad 0 \leq \xi \leq t_o \ ,$$

$j = 1, 2, 3, \ldots$. From hypothesis (3.1.6), and relations (3.2.4), (3.2.5), we find that

$$(3.2.18) \quad \liminf_{j \to \infty} \mathcal{k}(\xi,x(\xi,u_{n_j}),\sigma_{n_j}(\xi)) \geq \mathcal{k}(\xi,x(\xi,u_o),u_o(\xi)),$$

a.e. on $[0,t_o]$. Utilizing (3.2.7), we find that

$$(3.2.19) \qquad \liminf_{j \to \infty}\int_0^{t_o}\mathcal{k}(\xi,x(\xi,u_{n_j}),\sigma_{n_j}(\xi))d\xi \leq$$

$$\liminf_{k \to \infty}\frac{1}{n_j}\sum_{i=1}^{n_j}\int_0^{t_o}\mathcal{k}(\xi,x(\xi,u_{n_j}),u_i(\xi))d\xi \ .$$

By virtue of (3.2.17) and (3.2.18), Fatou's lemma [XIII, pg. 167] can be applied to the left hand side of (3.2.19) to obtain the relation

(3.2.20) $\liminf_{j\to\infty} \int_0^{t_o} \ell(\xi, x(\xi, u_{n_j}), \sigma_{n_j}(\xi))d\xi \geq \int_0^{t_o} \ell(\xi, x(\xi, u_o), u_o(\xi))d\xi.$

In view of the definition of K (see (3.1.7)), relations (3.2.9), (3.2.19), and (3.2.20), we infer that

(3.2.21) $$K(u_o, t_o) \leq \gamma.$$

But as a consequence of the definition of γ and the fact that $(u_o, t_o) \varepsilon\ \hat{\mathcal{U}}(\Omega, \Gamma)$ we must also have

(3.2.21') $$K(u_o, t_o) \geq \gamma.$$

Combining (3.2.21) and (3.2.21') we deduce that $K(u_o, t_o) = \gamma$, and this completes the proof for Case 1.

Case 2 $\qquad 0 \leq t_n \leq t_o \leq T$, $n = 1, 2, 3, \ldots$.

We extend each control in the "minimizing subsequence" to the entire interval $[0, t_o]$ in the manner prescribed in Definition 3.1.1, viz., let $u^* \varepsilon \mathcal{A}_0^{t_o}(\Omega)$ $(\mathcal{A}_0^{t_o}(\Omega) \neq \emptyset$ by Theorem 2.2.3), then define

$$\bar{u}_n(t) = \begin{cases} u_n(t), & \text{if } 0 \leq t \leq t_n \\ u^*(t), & \text{if } t_n < t \leq t_o, \end{cases}$$

and $\bar{u}_n \varepsilon \mathcal{A}_0^{t_o}(\Omega)$, $n = 1, 2, 3, \ldots$. Case 2 can then be disposed of by retracing the steps used in proving *Case 1* utilizing the "modified minimizing subsequence" in lieu of the one used in the proof for *Case 1*.

A good many of the details in the proofs of Theorems 3.22 and 3.23 (below) are quite similar to those given in the proof of Theorem 3.2.1. For this reason we shall use an abbreviated exposition in the proofs of the next two theorems.

THEOREM 3.2.2. *Let the hypotheses of Theorem* 3.1.1 *and Hypothesis* 3.1.6 *remain in effect. In addition suppose the following hypotheses are also satisfied:*

(i) $\hat{\mathcal{U}}(\Omega,\Gamma) \neq \emptyset$;

(ii) *For each* $t\varepsilon[0,T]$, $\mathscr{k}$ *is convex in the variables* $(x,u)\varepsilon R^n\times R^m$, i.e., *for each* $t\varepsilon[0,T]$ *and* (x,u), $(y,v)\varepsilon R^n\times R^m$, p, $q,\geq 0$, $p+q = 1$ *the inequality*

$$\mathscr{k}(t,px+qy,pu+qv)\leq p\,\mathscr{k}(t,x,u)+q\,\mathscr{k}(t,y,v)$$

holds (the special case $p = q = \frac{1}{2}$ *suffices for the proof of this theorem). Then there is a* $(u_0,t_0)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$ *such that* $K(u_0,t_0) = \inf K(\hat{\mathcal{U}}(\Omega,\Gamma)) \underset{d}{\equiv} \gamma$.

Proof: We may assume there is a sequence

$$(u_n,t_n)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma),\ n = 1,\ 2,\ 3,\ \ldots \tag{3.2.22}$$

such that

(3.2.23) (a) $t_n \to t_o \varepsilon \Gamma$ (monotonely) as $n \to \infty$;

(b) $u_n \to u_o$(wk) as $n \to \infty$, $(u_o, t_o) \varepsilon \hat{\mathcal{U}}(\Omega, \Gamma)$;

(c) $K(u_n, t_n) \to \gamma > -\infty$ as $n \to \infty$.

Just as in the proof of Theorem 3.2.1, there are two cases:

Case I $\qquad 0 \leq t_o \leq t_n \leq T$, $n = 1, 2, 3, \ldots$;

Case II $\qquad 0 \leq t_n \leq t_o \leq T$, $n = 1, 2, 3, \ldots$.

We consider Case I. By the Banach-Saks theorem (Theorem 2.2.1) there is a subsequence of $\{(u_n, t_n)\}$, which we still denote by $\{(u_n, t_n)\}$ such that

$$(3.2.24) \qquad \sigma_n = \frac{1}{n} \sum_{i=1}^{n} u_i \to u_o(st), \; t_n \to t_o \quad \text{as} \quad n \to \infty.$$

Then (consult [XI, pg. 87]) there is a subsequence of $\{(\sigma_n, t_n)\}$ say $\{(\sigma_{n_j}, t_{n_j})\}$ such that

$$(3.2.25) \qquad \sigma_{n_j} \to u_o(a.e.) \quad \text{on} \quad [0, t_o], \; t_{n_j} \to t_o \quad \text{as} \quad j \to \infty .$$

It follows from (3.2.23) (b) and (3.1.4) that

$$(3.2.26) \qquad x(t, u_n) \to x(t, u_o) \quad \text{as} \quad n \to \infty, \; 0 \leq t \leq t_o .$$

From (3.2.26), (3.2.23) (b), (c) we obtain that

$$(3.2.27) \quad \text{(a)} \quad \lim_{j \to \infty} \frac{1}{n_j} \sum_{i=1}^{n_j} x(t, u_i) = x(t, u_o), \quad 0 \leq t \leq t_o;$$

$$\text{(b)} \quad \lim_{n \to \infty} K(u_n, t_n) = \gamma = \lim_{n \to \infty} \int_0^{t_o} k(\xi, x(\xi, u_n), u_n(\xi)) d\xi;$$

$$\text{(c)} \quad \lim_{j \to \infty} \frac{1}{n_j} \sum \int_0^{t_o} k(\xi, x(\xi, u_i), u_i(\xi)) d\xi = \gamma.$$

The convexity condition (hypothesis (ii)) on $\mathcal{k}$ gives

$$(3.2.28) \qquad \int_0^{t_0} \mathcal{k}(\xi, \frac{1}{n_j}\sum_{i=1}^{n_j} x(\xi, u_i), \frac{1}{n_j}\sum_{i=1}^{n_j} u_i(\xi))d\xi$$

$$\leq \frac{1}{n_j}\sum_{i=1}^{n_j}\int_0^{t_0} \mathcal{k}(\xi, x(\xi, u_i), u_i(\xi))d\xi \ .$$

Then in a manner entirely similar to that presented in the proof of Theorem (3.2.1) one may establish that

$$\gamma \leq K(u_0, t_0) \leq \gamma,$$

which completes the proof for Case I.

The details of the proof for Case II will be omitted.

The force of Theorem 3.2.2 is that the Lipschitz condition (hypothesis (iii) of Theorem 3.2.1) on $\mathcal{k}$ may be dispensed with, if we require that $\mathcal{k}$ be a convex function in both the variables (x,u). The question which now arises is: Can the Lipschitz condition (hypothesis (iii) Theorem 3.2.1) and the convexity requirement in the variable x both be omitted? Insofar as the author has been able to to discover, the answer is a qualified yes, but at the expense of imposing other very restrictive hypotheses on the function $\mathcal{k}$ such as those given in the following theorem.

THEOREM 3.2.3. Let the hypotheses of Theorem 3.1.1 be satisfied. Suppose the following conditions are also true:

(i) $\hat{\mathcal{U}}(\Omega,\Gamma) \neq \emptyset$;

(ii) $\mathcal{k}(t,x,u) = a(t,x)+b_i(t,x)u^i(t)+c_{ij}(t,x)u^i(t)u^j(t)$

(*with summation on* $i, j = 1, 2, \ldots, m$), *and the mappings* a, b_i, c_{ij}, $i, j = 1, 2, \ldots, m$ *are required to satisfy the additional restrictions*:

(α) $a, b_i, c_{ij} : [0,T] \times R^n \to R$, $i, j = 1, 2, \ldots, m$ *are each continuous*;

(β) *For each* $(t,x,u) \varepsilon [0,T] \times R^n \times R^m$ *we have* $c_{ij}(t,x)u^i u^j \geq 0$ (*with summation on* $i, j = 1, 2, \ldots, m$) *and* $c_{ij} = c_{ji}$, $i, j = 1, 2, \ldots, m$.

Then there is a $(u_o, t_o) \varepsilon \hat{\mathcal{U}}(\Omega,\Gamma)$ *such that* $K(u_o,t_o) = \inf K(\hat{\mathcal{U}}(\Omega,\Gamma)) \equiv \gamma$

Proof: In accordance with Theorem 3.1.1 there is a "minimizing sequence" $\{(u_n,t_n)\}$ in $\hat{\mathcal{U}}(\Omega,\Gamma)$ such that

(3.2.29) (a) $u_n \to u_o(wk)$, $t_n \to t_o$ (monotonely) as $n \to \infty$, and

$(u_o,t_o) \varepsilon \hat{\mathcal{U}}(\Omega,\Gamma)$;

(b) $K(u_n,t_n) \to \gamma > -\infty$ as $n \to \infty$.

As in the proofs of the two preceding theorems, we consider the two cases:

Case 1 $\quad 0 \leq t_o \leq t_n \leq T$, $n = 1, 2, 3, \ldots$;

Case 2 $\quad 0 \leq t_n \leq t_o \leq T$, $n = 1, 2, 3, \ldots$.

First assume that *Case 1* holds. Then from (3.2.29) (a) and (3.1.4) it follows that

(3.2.30) $x(t,u_n) \to x(t,u_o)$ as $n \to \infty$ on $[0,t_o]$.

By hypothesis (ii) (α) of this theorem, relation (3.2.30), and the Lebesgue dominated convergence theorem (its applicability is easily justified), it is determined that

$$(3.2.31) \qquad \lim_{n\to\infty} \int_0^{t_o} a(\xi, x(\xi, u_n))d\xi = \int_0^{t_o} a(\xi, x(\xi, u_o))d\xi \quad .$$

Also by (3.2.30) and hypothesis (ii) (α) of this theorem we deduce that

$$(3.2.32) \qquad \lim_{n\to\infty} b_i(t, x(t, u_n)) = b_i(t, x(t, u_o)),$$

$$i = 1, 2, \ldots, m, \text{ whenever } 0 \leq t \leq t_o \quad .$$

The following inequality is easily seen to be valid:

$$(3.2.33) \qquad \left| \int_0^{t_o} \{b_i(\xi, x(\xi, u_n))u_n^i(\xi) - b_i(\xi, x(\xi, u_o))u_o^i(\xi)\}d\xi \right| \leq$$

$$\left| \int_0^{t_o} \{b_i(\xi, x(\xi, u_n)) - b_i(\xi, x(\xi, u_o))\}u_n^i(\xi)d\xi \right| +$$

$$\left| \int_0^{t_o} b_i(\xi, x(\xi, u_o))\{u_n^i(\xi) - u_o^i(\xi)\}d\xi \right| ,$$

for $n = 1, 2, 3, \ldots$ (summation on $i = 1, 2, \ldots, m$). Both terms on the right hand side of inequality (3.2.33) converge to zero as $n\to\infty$. The first term does so by equation (3.2.32), the fact that $\{u_n^i(\xi) |\ i = 1, 2, \ldots, m;\ n = 1, 2, 3, \ldots,\ 0\leq\xi\leq t_o\}$ is a bounded set of real numbers, and the Lebesgue dominated convergence theorem. The second term on the right hand side of inequality (3.2.33) converges to zero as $n\to\infty$ because of the weak convergence of the sequence $\{u_n - u_o\}$ to zero as $n\to\infty$. We thereby have that

$$(3.2.34) \quad \lim_{n\to\infty} \int_0^{t_o} b_i(\xi, x(\xi, u_n))u_n^i(\xi)d\xi = \int_0^{t_o} b_i(\xi, x(\xi, u_o))u_o^i(\xi)d\xi$$

(with summation on $i = 1, 2, \ldots, m$). Next it will be established that

$$(3.2.35)\qquad \liminf_{n\to\infty} \int_0^{t_o} c_{ij}(\xi,x(\xi,u_n))u_n^i(\xi)u_n^j(\xi)d\xi \geq \int_0^{t_o} c_{ij}(\xi,x(\xi,u_o))u_o^i(\xi)u_o^j(\xi)d\xi .$$

Define $s_n = u_n - u_o$, $n = 1, 2, 3, \ldots,$ then $s_n \to 0(wk)$ as $n \to \infty$, and $u_n = s_n + u_o \to u_o(wk)$ as $n \to \infty$. We have the equality

$$(3.2.36)\qquad c_{ij}(\xi,x(\xi,u_n))u_n^i(\xi)u_n^j(\xi) = c_{ij}(\xi,x(\xi,u_n))s_n^i(\xi)s_n^j(\xi) + c_{ij}(\xi,x(\xi,u_n))u_o^i(\xi)s_n^j(\xi) + c_{ij}(\xi,x(\xi,u_n))s_n^i(\xi)u_o^j(\xi) + c_{ij}(\xi,x(\xi,u_n))u_o^i(\xi)u_o^j(\xi) ,$$

for $0 \leq \xi \leq t_o$. It follows from (3.2.36) and hypothesis (ii) (β) of this theorem that

$$(3.2.37)\qquad c_{ij}(\xi,x(\xi,u_n))u_n^i(\xi)u_n^j(\xi) \geq 2c_{ij}(\xi,x(\xi,u_n))s_n^i(\xi)u_o^j(\xi) + c_{ij}(\xi,x(\xi,u_n))u_o^i(\xi)u_o^j(\xi)$$

for $0 \leq \xi \leq t_o$. From (3.2.37) it follows that

$$(3.2.38)\qquad \int_0^{t_o} c_{ij}(\xi,x(\xi,u_n))u_n^i(\xi)u_n^j(\xi)d\xi \geq 2\int_0^{t_o} c_{ij}(\xi,x(\xi,u_n))s_n^i(\xi)u_o^j(\xi)d\xi + \int_0^{t_o} c_{ij}(\xi,x(\xi,u_n))u_o^i(\xi)u_o^j(\xi)d\xi .$$

It is a simple matter to prove that

$$(3.2.39) \qquad \lim_{n\to\infty} 2\int_0^{t_o} c_{ij}(\xi,x(\xi,u_n))s_n^i(\xi)u_o^j(\xi)d\xi \; +$$

$$\lim_{n\to\infty} \int_0^{t_o} c_{ij}(\xi,x(\xi,u_n))u_o^i(\xi)u_o^j(\xi)d\xi \; =$$

$$\int_0^{t_o} c_{ij}(\xi,x(\xi,u_o))u_o^i(\xi)u_o^j(\xi)d\xi \; .$$

In order to demonstrate the validity of (3.2.39), we observe that because of (3.2.30), and the continuity of the c_{ij}, $i, j = 1, 2, \ldots, m$, the second term on the left hand side of equation (3.2.39) coincides with the term on the right hand side of the same equation. It therefore clearly suffices to prove

$$(3.2.40) \qquad \lim_{n\to\infty} 2\int_0^{t_o} c_{ij}(\xi,x(\xi,u_n))s_n^i(\xi)u_o^j(\xi)d\xi =$$

$$\lim_{n\to\infty} 2\int_0^{t_o} c_{ij}(\xi,x(\xi,u_o))s_n^i(\xi)u_o^j(\xi)d\xi \; ,$$

since the right hand side of (3.2.40) is clearly zero in view of the fact that $s_n \to 0(wk)$ as $n\to\infty$. The $\{s_n^i\} = \{u_n^i - u_o^i\}$ are uniformly bounded on $[0,t_o]$. Thus in view of (3.2.30) and the continuity of the c_{ij}, $i, j = 1, 2, \ldots, m$ we have that

$$\lim_{n\to\infty} \int_0^{t_o} \{c_{ij}(\xi,x(\xi,u_n)) - c_{ij}(\xi,x(\xi,u_o))\}s_n^i(\xi)u_o^j(\xi)d\xi \; = 0,$$

thereby proving (3.2.40). By (3.2.38) and (3.2.39) we have that

$$(3.2.41) \qquad \liminf_{n\to\infty} \int_0^{t_o} c_{ij}(\xi,x(\xi,u_n))u_n^i(\xi)u_n^j(\xi)d\xi \geq$$

$$\int_0^{t_o} c_{ij}(\xi,x(\xi,u_o))u_o^i(\xi)u_o^j(\xi)d\xi \; .$$

It is once again trivial to prove that

$$(3.2.42) \quad \underline{\lim}_{n\to\infty} K(u_n,t_n) = \underline{\lim}_{n\to\infty} \int_0^{t_o} k(\xi,x(\xi,u_n),u_n(\xi))\, d\xi$$

where k is given in hypothesis (ii) of this theorem. Therefore by (3.2.31), (3.2.34), (3.2.41), and (3.2.42) we obtain that

$$\gamma = \liminf_{n\to\infty} K(u_n,t_n) \geq K(u_o,t_o) \geq \gamma,$$

with $(u_o,t_o) \in \hat{\mathcal{U}}(\Omega,\Gamma)$. Thus the proof for Case 1 is complete, since $K(u_o,t_o) = \gamma$.

We omit the details of the proof for Case 2.

3.3 Applications

Let $(R^n, ||\cdot||)$ be the usual euclidean space of n-dimensions, where for $x = (x^1,x^2,\ldots,x^n) \in R^n, ||x||^2 = \sum (x^i)^2$. For $\varepsilon>0$ define

$$J_\varepsilon = \{(x,y) | x,y \in R^n, ||x-y|| < \varepsilon\}.$$

Then $\mathcal{J} = [\{J_\varepsilon | \varepsilon>0\}]$ is the uniform structure on R^n induced by the norm $||\cdot||$. Also for $A \subset R^n$ we define the "diameter of A" to be

$$(3.3.1) \qquad \delta(A) = \sup \{||x-y|| \ | x, y \in A\} .$$

We also shall have occasion to use

$$(3.3.2) \qquad \lambda(A,B) = \inf \{||x-y|| \ | x \in A, y \in B\} ,$$

$A,B \subset R^n$.

Let $[a,b] \subset R$ be a compact interval. Suppose F is a mapping, $F:[a,b] \to \mathcal{C}(R^n)$ (for the meaning of the notation consult Chapter II). If $F(t) \neq \emptyset$, $a \leq t \leq b$, then we define a mapping

$\mathcal{k}:[a,b]\times R^n \to R$ by the equation

$$\mathcal{k}(t,x) = \lambda(\{x\}, F(t)). \tag{3.3.3}$$

where λ is defined by (3.3.2). We have the following theorem.

THEOREM 3.3.1. <u>Suppose</u> F <u>is usci on</u> $[a,b]$, <u>then</u> $\mathcal{k}:[a,b]\times R^n \to R$ <u>is lsc on</u> $[a,b]\times R^n$.

Proof: It is well known that for fixed $t\varepsilon[a,b]$, $\mathcal{k}(t,\cdot):R^n \to R$ is continuous [IX, pg. 120]. It will now be demonstrated that $\mathcal{k}(\cdot,x)$ is lsc on $[a,b]$ (uniformly with respect to $x\varepsilon R^n$). Since F is usci on $[a,b]$, then for $t_o\varepsilon[a,b]$, $\varepsilon>0$, there is a $\delta_{\varepsilon,t_o}>0$ such that

$$t\varepsilon[a,b], |t-t_o|<\delta_{\varepsilon,t_o} \text{ imply } F(t)\subset J_{\varepsilon/2}[F(t_o)]. \tag{3.3.4}$$

Thus if $t\varepsilon[a,b], |t-t_o|<\delta_{\varepsilon,t_o}$, there is a $y_{t,x}\varepsilon F(t)$ such that $\mathcal{k}(t,x)+\varepsilon/2 > ||x-y_{t,x}||$, and by (3.3.4) there is a $b_{t_o}\varepsilon F(t_o)$ such that $||y_{t,x}-b_{t_o}||<\varepsilon/2$. Now by the definition of $\mathcal{k}(t_o,x)$ (equation (3.3.3)) we have

$$\mathcal{k}(t_o,x)\leq||x-b_{t_o}|| \leq ||x-y_{t,x}|| + ||y_{t,x}-b_{t_o}||.$$

Thus if $t\varepsilon[a,b]$, $|t-t_o|<\delta_{\varepsilon,t_o}$, then

$$\mathcal{k}(t_o,x)<\mathcal{k}(t,x)+\varepsilon/2+\varepsilon/2 = \mathcal{k}(t,x)+\varepsilon$$

Whence $\mathcal{k}(\cdot,x)$ is lsc at $t_o\varepsilon[a,b]$, and the above $\delta_{\varepsilon,t_o}>0$

depends only on ε and t_o. Finally it must be established that $\mathcal{k}:[a,b]\times R^n \to R$ is lsc. Thus given $t_o \varepsilon [a,b]$ and $\varepsilon>0$ pick $\delta_{\varepsilon,t_o}>0$ such that

(3.3.5) $\quad t\varepsilon[a,b], |t-t_o|<\delta_{\varepsilon,t_o}$ imply $\mathcal{k}(t,x)>\mathcal{k}(t_o,x)-\varepsilon/2,$

for $x\varepsilon R^n$. Let $x_o \varepsilon R^n$, then because $\mathcal{k}(t_o,\cdot)$ is continuous at x_o, there is a $\delta^*_{t_o,x_o,\varepsilon}>0$ such that

(3.3.6) $\quad ||x-x_o|| < \delta^*_{t_o,x_o,\varepsilon}$ implies $|\mathcal{k}(t_o,x)-\mathcal{k}(t_o,x_o)| < \varepsilon/2.$

Define $\delta_1 = \min(\delta_{\varepsilon,t_o}, \delta^*_{t_o,x_o,\varepsilon})$, then $||x-x_o||<\delta_1$ and $|t-t_o|<\delta_1$, $t\varepsilon[a,b]$ imply

(3.3.7) $\qquad$ (a) $\mathcal{k}(t,x)>\mathcal{k}(t_o,x)-\varepsilon/2,$

$\qquad$ (b) $\mathcal{k}(t_o,x)>\mathcal{k}(t_o,x_o)-\varepsilon/2.$

Combining (3.3.7) (a) and (b) we obtain that $\mathcal{k}$ is lsc at $(t_o,x_o)\varepsilon[a,b]\times R^n$.

THEOREM 3.3.32. <u>If</u> F <u>is lsci on</u> $[a,b]$, <u>then</u> $\mathcal{k}$ <u>is usc on</u> $[a,b]\times R^n$.

Proof: Just as in the proof of Theorem 3.3.1, the mapping $\mathcal{k}(t,\cdot)$ from R^n to R is continuous (t is fixed, $a\le t\le b$). It will now be proved that $\mathcal{k}(\cdot,x):[a,b]\to R$ is usc (uniformly with respect to $x\varepsilon R^n$). Suppose $t_o\varepsilon[a,b]$ and $\varepsilon>0$, then in view of the assumption that F is lsci on $[a,b]$, (in particular at t_o) it follows that there is a $\delta_{\varepsilon,t_o}>0$ such that

(3.3.8) $\quad t\epsilon[a,b], |t-t_o|<\delta_{\epsilon,t_o}$ imply $J_{\epsilon/2}[F(t)] \supset F(t_o)$.

By the definition of $\mathcal{k}(t_o,x)$, there is a $y_{t_o,x}\epsilon F(t_o)$ such that

(3.3.9) $$||x-y_{t_o,x}|| < \mathcal{k}(t_o,x)+\epsilon/2 .$$

From (3.3.8) we obtain that if $t\epsilon[a,b], |t-t_o|<\delta_{\epsilon,t_o}$, then there is a $b_t\epsilon F(t)$ such that $||b_t-y_{t_o,x}||<\epsilon/2$. Whence if $t\epsilon[a,b], |t-t_o|<\delta_{\epsilon,t_o}$, then from 3.3.9 it follows that

(3.3.10) $$\mathcal{k}(t,x)\leq||x-b_t||\leq||x-y_{t_o,x}||+||y_{t_o,x}-b_t|| < \mathcal{k}(t_o,x)+\epsilon/2+\epsilon/2 = \mathcal{k}(t_o,x)+\epsilon .$$

It follows from (3.3.10) that $\mathcal{k}(\cdot,x)$ is usc at $t_o\epsilon[a,b]$ (uniformly with respect to $x\epsilon R^n$). Thus if $(t_o,x_o)\epsilon[a,b]\times R^n$ and $\epsilon>0$, there is a $\delta_{\epsilon,t_o}>0$ (depending only on $\epsilon>0$ and $t_o\epsilon[a,b]$) such that

(3.3.11) $\quad t\epsilon[a,b], |t-t_o|<\delta_{\epsilon,t_o}$ imply $\mathcal{k}(t,x)<\mathcal{k}(t_o,x)+\epsilon/2$,

for $x\epsilon R^n$. Because $\mathcal{k}(t_o,\cdot)$ is continuous at $x_o\epsilon R^n$ there is a $\delta^*_{t_o,\epsilon,x_o}>0$ such that

(3.3.12) $\quad ||x-x_o||<\delta^*_{t_o,\epsilon,x_o}$ implies $|\mathcal{k}(t_o,x_o)-\mathcal{k}(t_o,x)|<\epsilon/2$.

Define $\delta_1 = \min(\delta^*_{t_o,\varepsilon,x_o}, \delta_{\varepsilon,t_o})$, then combining inequalities (3.3.11), (3.2.12) we obtain that

(3.3.13) $\quad t\varepsilon[a,b], |t-t_o|<\delta_1, ||x-x_o||<\delta_1$ imply

$$\mathcal{k}(t,x) < \mathcal{k}(t_o,x_o)+\varepsilon \ .$$

Whence $\mathcal{k}$ is usc at $(t_o,x_o)\varepsilon[a,b]\times R^n$. This completes the proof.

"Partial" converses of Theorems 3.3.1 and 3.3.2 are also true, viz.,

THEOREM 3.3.3. <u>If</u> $\mathcal{k}(\cdot,x):[a,b]\to R$ <u>is lsc (uniformly with respect to</u> $x\varepsilon R^n$), <u>then</u> F <u>is usci on</u> [a,b].

Proof: Suppose $t_o\varepsilon[a,b], \varepsilon>0$, then there is a $\delta_{\varepsilon,t_o}>0$ such that

(3.3.14) $\quad t\varepsilon[a,b], |t-t_o|<\delta_{\varepsilon,t_o}$ imply $\mathcal{k}(t,x) > \mathcal{k}(t_o,x)-\varepsilon$,

for $x\varepsilon R^n$. Thus if $t\varepsilon[a,b]$ and $|t-t_o|<\delta_{\varepsilon,t_o}$, and $x\varepsilon F(t)$, then $\mathcal{k}(t,x) = \inf\{||x-y|| \mid y\varepsilon F(t)\} = 0$. Whence $0 \leq \mathcal{k}(t_o,x) < \varepsilon$. Consequently there is a $y\varepsilon F(t_o)$ such that $||x-y||<\varepsilon$ (by the definition of $\mathcal{k}(t_o,x)$). Therefore $x\varepsilon J_\varepsilon[F(t_o)]$, and we have thereby proved

$$t\varepsilon[a,b], |t-t_o|<\delta_{\varepsilon,t_o} \text{ imply } F(t) \subset J_\varepsilon[F(t_o)] \ .$$

This completes the proof.

We also obtain

THEOREM 3.3.4. *If* $k(\cdot,x):[a,b]\to R$ *is usc on* $[a,b]$ *(uniformly with respect to* $x\varepsilon R^n$*), then* F *is lsci.*

Proof: Let $t_o\varepsilon[a,b]$, and suppose $\varepsilon>0$. Then there is a $\delta_{\varepsilon,t_o}>0$ such that

$$(3.3.15)\qquad t\varepsilon[a,b], |t-t_o|<\delta_{\varepsilon,t_o} \text{ imply } k(t,x)<k(t_o,x)+\varepsilon,$$

for $x\varepsilon R^n$. Let $t\varepsilon[a,b], |t-t_o|<\delta_{\varepsilon,t_o}$, $x\varepsilon F(t_o)$, then $k(t_o,x)=0$. Whence $0\leq k(t,x)<\varepsilon$. Thus from the definition of $k(t,x)$ there must exist a $y\varepsilon F(t)$ such that $||x-y||<\varepsilon$. Therefore $x\varepsilon J_\varepsilon[F(t)]$. In summary, it has been demonstrated that

$$(3.3.16)\qquad t\varepsilon[a,b], |t-t_o|<\delta_{\varepsilon,t_o} \text{ imply } J_\varepsilon[F(t)]\supset F(t_o),$$

which proves that F is lsci at $t_o\varepsilon[a,b]$.

COROLLARY 3.3.1. *If* F *is continuous on* $[a,b]$*, then* k *is continuous on* $[a,b]\times R^n$*. Moreover, if* $k(\cdot,x)$ *is continuous (uniformly with respect to* $x\varepsilon R^n$*) on* $[a,b]$*, then* F *is continuous on* $[a,b]$.

Proof: When F is continuous on $[a,b]$, then F is both usci and lsci on $[a,b]$ (Theorem 2.1.2). Thus by Theorems 3.3.1 and 3.3.2 k is both lsc and usc on $[a,b]\times R^n$, and consequently k is continuous on $[a,b]\times R^n$. On the other hand if $k(\cdot,x)$ is continuous (uniformly with respect to $x\varepsilon R^n$) on $[a,b]$, then the hypotheses of Theorems 3.3.3 and 3.3.4 are fulfilled. Therefore F is both usci and lsci on $[a,b]$, and so by Theorem 2.2.2 F must be continuous on $[a,b]$.

THEOREM 3.3.5. *If* $F(t)$ *is convex for each* $t \in [a,b]$, *then* $\mathscr{k}$ *is a convex function in the variable* $x \in R^n$.

Proof: See [II, pg. 5].

Remark 3.3.1: We *cannot* conclude from the hypothesis of Theorem 3.3.1 that $\mathscr{k}$ is continuous. Consider the example: Let $F:[0,1] \to \mathscr{C}(R^2)$ be defined by

$$F(t) = \begin{cases} \{(x,y) \mid x^2+y^2 \leq 1,\ x,\ y \in R\}, & \text{if } t = 0 \\ \{(x,y) \mid (x-1)^2+y^2 \leq t^2,\ x,\ y \in R\}, & \text{if } t \neq 0 . \end{cases}$$

Suppose $h = (1,2) \in R^2$, then

$$\mathscr{k}(t,h) = \begin{cases} \sqrt{6-2\sqrt{5}} \neq 2, & \text{if } t = 0 \\ 2-t, & \text{if } 0 < t \leq 1 \end{cases} .$$

It is clear that F is usci on $[0,1]$, however, $\mathscr{k}(\cdot,h)$ is not continuous at $t = 0$, so $\mathscr{k}$ is not continuous. It can be shown that $\mathscr{k}$ is lsc on $[0,1] \times R^2$.

Remark 3.3.2: We *cannot* conclude from the hypothesis of Theorem 3.3.2 that $\mathscr{k}$ is continuous. Consider the example

$$F(t) = \begin{cases} \{(x,y) \mid x^2+y^2 \leq t^2,\ x,\ y \in R\}, & 0 \leq t < 1 \\ \{(0,0)\}, & t = 1 \end{cases} .$$

Let $h = (0,1)\varepsilon R^2$, then

$$\mathcal{k}(t,h) = \begin{cases} 1-t, & 0\leqslant t<1 \\ 1, & t = 1 \end{cases} .$$

Clearly $\mathcal{k}(\cdot,h)$ is not continuous at $t = 1$, and F is lsci on $[0,1]$. Evidently $\mathcal{k}$ is usc on $[0,1]\times R^2$.

Although we do not intend to make immediate use of the following two theorems, they appear to be of some interest, and are included for future reference.

THEOREM 3.3.6. <u>Let</u> F <u>be usci on</u> $[a,b]$. <u>Then the mapping</u> $\psi:[a,b]\to R^*::t\mapsto\psi(t) = \delta(F(t))$ <u>is usc on</u> $[a,b]$ (R^* <u>denotes the extended real number system; the definition of</u> δ <u>is found in</u> eq. (3.3.1)).

Proof: Let $t_o\varepsilon[a,b]$. It is clear from the definition of ψ that $0\leqslant\psi(t_o)\leqslant+\infty$. There are two cases:

<u>Case 1</u> $\psi(t_o) = +\infty$. The proof is trivial in this case.

<u>Case 2</u> $0\leqslant\psi(t_o)<+\infty$.

Given $\varepsilon>0$, there is a $\delta_{\varepsilon,t_o}>0$ such that

(3.3.17) $t\varepsilon[a,b],|t-t_o|<\delta_{\varepsilon,t_o}$ imply $J_{\varepsilon/3}[F(t_o)]\supset F(t)$.

It will now be established that

(3.3.18) $$t\varepsilon[a,b],|t-t_o|<\delta_{\varepsilon,t_o} \text{ imply}$$

$$\delta(F_t)\leqslant\delta(J_{\varepsilon/3}[(F_t)])<\delta(F_{t_o})+\varepsilon ,$$

The first part of the inequality in (3.3.18) is clear, in view of (3.3.17). Thus we shall only prove that $(J_{\varepsilon/3}[F(t_o)]) < \delta(F(t_o))+\varepsilon$. Let x, $y\varepsilon J_{\varepsilon/3}[F(t_o)]$, then there exist a_o, $b_o \varepsilon F(t_o)$ such that

(3.3.19) $$||x-a_o|| < \varepsilon;\ ||y-b_o|| < \varepsilon/3.$$

But since a_o, $b_o \varepsilon F(t_o)$, it follows from the definition of δ that

(3.3.20) $$||a_o-b_o|| < \delta(F(t_o)).$$

Combining inequalities (3.3.19) and (3.3.20) we conclude that

(3.3.21) $$\begin{aligned}||x-y|| &\leq ||x-a_o|| + ||a_o-b_o||+||b_o-y||\\ &< \varepsilon/3 + \varepsilon/3 + \delta(F(t_o)) = \frac{2\varepsilon}{3} + \delta(F(t_o)).\end{aligned}$$

Since x,y were arbitrary elements of $J_{\varepsilon/3}[F(t_o)]$, it follows from the definition of δ that

$$\delta(J_{\varepsilon/3}[F(t_o)]) < \frac{2}{3}\varepsilon + \delta(F(t_o)) < \delta(F(t_o)) + \varepsilon.$$

The validity of (3.3.18) is thereby established, and consequently ψ is usc at $t_o \varepsilon [a,b]$.

THEOREM 3.3.7. <u>If</u> F <u>is lsci on</u> $[a,b]$ <u>then the mapping</u> ψ <u>defined in Theorem</u> 3.36 <u>is lsc on</u> $[a,b]$.

Proof: Suppose $t_o \varepsilon [a,b]$, and $\varepsilon > 0$. Then there is a $\delta_{\varepsilon,t_o} > 0$ such that

(3.3.22) $$t\varepsilon [a,b], |t-t_o| < \delta_{\varepsilon,t_o} \text{ imply } J_{\varepsilon/3}[F(t)] \supset F(t_o).$$

From (3.3.22) it follows that

(3.3.23) $t\varepsilon[a,b], |t-t_o| < \delta_{\varepsilon,t_o}$ imply $\delta(F(t_o)) \leq \delta(J_{\varepsilon/3}[F(t)])$.

If $\delta(F(t_o)) = +\infty$, then $\delta(J_{\varepsilon/3}[F(t)]) = +\infty$, for $t\varepsilon[a,b]$, $|t-t_o| < \delta_{\varepsilon,t_o}$. Thus let $h\varepsilon R$, then there exist $x_h, y_h \varepsilon J_{\varepsilon/3}[F(t)]$ such that

(3.3.24) $$||x_h - y_h|| > h.$$

But since $x_h, y_h \varepsilon J_{\varepsilon/3}[F(t)]$, there must exist $a,b\varepsilon F(t)$ such that

(3.3.25) $$||x_h - a|| < \varepsilon/3, \quad ||y_h - b|| < \varepsilon/3.$$

Thus by (3.3.24) and (3.3.25)

(3.3.26) $$h < ||x_h - y_h|| \leq ||x_h - a|| + ||a-b|| + ||b-y_h|| < ||a-b|| + \frac{2\varepsilon}{3}.$$

Whence $t\varepsilon[a,b], |t-t_o| < \delta_{\varepsilon,t_o}$ imply

$$\delta(F(t)) \geq h - \frac{2\varepsilon}{3}.$$

Therefore

$$\inf\ \{\delta(F(t)) \mid |t-t_o| < \delta_{\varepsilon,t_o},\ t\varepsilon[a,b]\} > h - \frac{2\varepsilon}{3},$$

and consequently

$$\liminf_{t\to t_o}\ \delta(F(t)) \geq h - \frac{2\varepsilon}{3}.$$

Since h was arbitrary, it follows that $\liminf_{t \to t_0} \delta(F(t)) = +\infty = \delta(F(t_0))$. Whence if $\delta(F(t_0)) = +\infty$, then ψ is lsc at t_0. If $\delta(F(t_0)) < +\infty$, then it will be demonstrated that

$$\delta(J_{\varepsilon/3}[F(t)]) \leq \delta(F(t)) + \varepsilon , \tag{3.3.27}$$

with equality holding iff $\delta(F(t)) = +\infty$. In case $\delta(F(t)) = +\infty$, the proof of (3.3.27) (with strict inequality) is very similar to the proof of (3.3.18) in Theorem 3.3.6, so we will not prove (3.3.27) in this case. Moreover, if $\delta(F(t)) = +\infty$, then both sides of (3.3.27) are $+\infty$. Thus using (3.3.23) and (3.3.27), and the assumption that $\delta(F(t_0)) < +\infty$, we have that

$$t \in [a,b], |t-t_0| < \delta_{\varepsilon,t_0} \text{ imply } \delta(F(t_0)) < \delta(F(t)) + \varepsilon . \tag{3.3.28}$$

In either case ($\delta(F(t_0)$ finite or infinite) ψ is lsc at t_0 and this completes the proof.

We are now able to generalize (in a sense which will be made more precise below) a problem stated by Pontryagin et al [XV, pg. 197 ff.] concerned with the application of the mathematical theory of optimal control to a problem in the approximation of functions. In this connection consider the linear optimal control problem (consult Section 3.1 for the formulation), with the "moving target set" $F:[0,T] \to \mathcal{C}(R^n)$ satisfying the hypotheses of Theorem 3.1.1 as well as the condition that $F(t)$ is convex for each $t \in [0,T]$, and $\bigcup_{0 \leq t \leq T} F(t)$ is bounded. In addition let all of the hypotheses of Theorem 3.1.1 remain in effect. Then

$$\beta = \{x(t,u) \mid (u,t_1) \varepsilon \mathcal{U}(\Omega,\Gamma),\ \ 0 \leq t \leq t_1\}\ \ ,$$

is a bounded subset of R^n (see equation 3.1.4). It may easily be demonstrated (in view of Theorem 3.3.1) that the mapping $\mathcal{k}:[0,T]\times co\beta \rightarrow R$ (where $co\beta$ denotes the convex hull of the set β) satisfies the conditions in hypothesis (3.1.6). Now instead of using F to define $\hat{\mathcal{U}}(\Omega,\Gamma)$ as was done in (3.1.5), we use $F^*[0,T] \rightarrow \mathcal{C}(R^n)::t \rightarrow F^*(t) = R^n$ to define $\hat{\mathcal{U}}(\Omega,\Gamma)$ in (3.1.5). With this understanding, it is clear that $\hat{\mathcal{U}}(\Omega,\Gamma) \neq \emptyset$. We infer from Theorem 3.3.5 that $\mathcal{k}$ satisfies hypothesis (ii) of Theorem 3.2.2. Thus if $\mathcal{h}:[0,T]\times R^n \times R^m \rightarrow R$ satisfies (3.1.6) and hypothesis (ii) of Theorem (3.2.2), then $\mathcal{k}+\mathcal{h}$ is a mapping of the same type satisfying (3.1.6) and hypothesis (ii) of Theorem 3.2.2. We therefore define $K: \hat{\mathcal{U}}(\Omega,\Gamma) \rightarrow R$ (note that F^* rather than F was used to define $\hat{\mathcal{U}}(\Omega,\Gamma)$) by the equation

$$(3.3.29)\qquad K(u,t_1) = \int_0^{t_1} \mathcal{k}(\xi,x(\xi,u)) + \mathcal{h}(\xi,x(\xi,u),u(\xi))d\xi,$$

for $(u,t_1)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$. Theorem 3.2.2 may then be applied to give the existence of a $(u_o,t_o)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$ such that

$$K(u_o,t_o) = \inf K(\hat{\mathcal{U}}(\Omega,\Gamma)) > -\infty.$$

Actually it appears that especially $\mathcal{h}(t,x,u) = ||u||^2$ in equation (3.3.29) would be a reasonable choice for a cost functional in many instances, and as was just pointed out some of our existence theory in this chapter applies to this case. The reader will no doubt perceive several other variations of this cost functional to which our results may be applied. The problem (mentioned above [XV, pg. 197 ff.]) which Pontryagin and his associates considered was the special case when $F(t) = \{y(t)\}$ = singleton point set, where, for example, $y:[0,T]\to R$ is required to be continuous and the cost functional is

$$\int_0^T (x(t)-y(t))^2 \, dt.$$

This functional is to be minimized on the class of all functions $x:[0,T]\to R$ which have continuous derivatives up to and including those of order n, and the n^{th} derivative $x^{(n)}$ satisfies a Lipschitz condition with constant α.

CHAPTER IV

THE EXISTENCE OF OPTIMAL CONTROLS

FOR

NONLINEAR SYSTEMS

4.1 Formulation of the Problem

In this chapter it will be assumed that the control system can be described by a system of real ordinary differential equations of the form

$$(4.1.1) \qquad \dot{x}^i = f^i(t,x^1,\ldots,x^n,u^1,\ldots,u^m), \quad i = 1, 2, \ldots, n,$$

where $\dot{x}$ as usual denotes differentiation with respect to the real independent variable t. If we write $f = (f^1,\ldots,f^n)$, $x = (x^1,\ldots,x^n)$, $u = (u^1,\ldots,u^m)$, then (4.1.1) can be more concisely written as

$$(4.1.1') \qquad \dot{x} = f(t,x,u).$$

The symbol u appearing in (4.1.1') is called the control parameter. As is customary when u is given as a function of t, u will be called a <u>control function</u> (or simply a control). The variables (components of u), u^i, $i = 1, \ldots, m$ will sometimes be referred to as control variables. In this chapter Ω is a <u>fixed</u> compact subset of R^m (Ω is nonempty), and $T > 0$ is a fixed real number. Suppose Γ

is a nonempty closed subset of $[0,T]$, then we define the set $\mathcal{U}(\Omega,\Gamma)$ as follows: $(u,t_1)\varepsilon\mathcal{U}(\Omega,\Gamma)$ iff

(i) u is a function, $u:[0,t_1]\to R^m$, and $u(t)\varepsilon\Omega$, $0 \leq t \leq t_1$ where $t_1\varepsilon\Gamma$;

(ii) $u\varepsilon L_2^m[0,t_1]$.

It will be necessary to make suitable assumptions on the functions f^i, $i = 1, 2, \ldots, n$ appearing in (4.1.1), which will guarantee that given $(u,t_1)\varepsilon\mathcal{U}(\Omega,\Gamma)$, $x_o\varepsilon R^n$, then there exists a unique absolutely continuous function (absolutely continuous <u>responses to the control</u> u) $x(\cdot,u):[0,t_1]\to R^n$ satisfying (4.1.1') almost everywhere (a.e.) on $[0,t_1]$ and the initial condition

$$x(0,u) = x_o. \tag{4.1.2}$$

We intend to make use of a classical uniqueness [XIII, pp. 345-346] and existence theorem "in the large" [XIII, pg. 342] in the theory of real ordinary differential equations. With this in mind the following assumptions are made:

(4.1.3) The mappings $f^i:R\times R^n\times R^m\to R$, $i = 1, 2, \ldots, n$ are continuous in (x,u) and integrable with respect to t for each fixed (x,u) in $R^n\times R^m$;

(4.1.4) The Lipschitz condition holds. There exists a constant $L > 0$ such that for any $(t,u)\varepsilon R\times R^m$

$$||f(t,x,u)-f(t,y,u)|| \leq L||x-x_1||, x,x_1\varepsilon R^n,$$

where for $x = (x^1,\ldots,x^n)$ R^n we define $||x||^2=\sum(x^i)^2$;

(4.1.5) For all $u \varepsilon R^m$ (uniformly)

$$||f(t,x,u)|| \leqslant \mu(t)[C+||x||]$$

where μ is integrable on every finite interval, and C is a positive constant.

Then hypotheses (4.1.3) through (4.1.5) are sufficient to insure the validity of the desired result: Given $(u,t_1)\varepsilon \mathcal{U}(\Omega,\Gamma)$, $x_0 \varepsilon R^n$, there is a unique absolutely continuous response $x(\cdot,u)$ to the control u satisfying (5.1.1') a.e. on $[0,t_1]$ and the initial condition (4.1.2).

Let $\mathcal{J}$ be a mapping, $\mathcal{J}:[0,T]\to \mathcal{C}(R^n)$; such that $\mathcal{J}(t) \neq \emptyset$ for each $t\varepsilon[0,T]$, and let $\mathcal{J}$ be usc on $[0,T]$ (see Section 2.1 for the definition of this terminology). Define $\hat{\mathcal{U}}(\Omega,\Gamma)$ to be that particular subset of $\mathcal{U}(\Omega,\Gamma)$ given by

$$(4.1.6) \qquad \hat{\mathcal{U}}(\Omega,\Gamma) = \{(u,t_1)\varepsilon \mathcal{U}(\Omega,\Gamma) | x(t_1,u)\varepsilon \mathcal{J}(t_1)\} .$$

We introduce the mapping $f^o:R\times R^n \times R^m \to R$ subject to the same restrictions which have been imposed on the f^i, $i = 1, \ldots, n$ in hypotheses (4.1.3) through (4.1.5). Then the mapping $K:\mathcal{U}(\Omega,\Gamma)\to R$ given by

$$(4.1.7) \qquad K(u,t_1) = \int_0^{t_1} f^o(\xi,x(\xi,u),u(\xi))d\xi ,$$

$(u_1,t_1)\varepsilon \mathcal{U}(\Omega,\Gamma)$, is well defined. The problem which we shall consider in this chapter is

$$(4.1.8) \qquad K(u,t_1) = \text{minimum on } \hat{\mathcal{U}}(\Omega,\Gamma).$$

As was suggested in the introductory chapter only one aspect of this problem, (4.1.8), will be discussed, viz., the existence of a $(u_o,t_o)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$ such that $K(u_o,t_o)$ is a minimum on $\hat{\mathcal{U}}(\Omega,\Gamma)$, i.e.,

$$K(u_o,t_o) \leq K(u,t_1), \quad \text{for each} \quad (u,t_1)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma).$$

Problem (4.1.8) will be called the <u>optimal control problem for nonlinear systems</u>.

It is our intention in this chapter to become in a certain sense more restrictive in our assumptions regarding $\hat{\mathcal{U}}(\Omega,\Gamma)$ than Roxin [17], but at the same time we shall relax considerably the requirements on the mappings f^i, $i = 0, 1, 2, \ldots, n$ in (4.1.1) and (4.1.7), and then demonstrate the validity of several existence theorems for problem (4.1.8). In particular Theorem 4.2.1 will yield as corollaries: an existence theorem proved by Lee and Markus [11, pp. 46-47], and an extension of the above cited theorem of Lee and Markus, as well as have important implications of its own.

<u>4.2 The Existence of a Solution to the Optimal Control Problem</u> (4.1.8) <u>within the Class</u> $\hat{\mathcal{U}}(\Omega,\Gamma)$.

DEFINITION 4.2.1. $\hat{\mathcal{U}}(\Omega,\Gamma)$ <u>is said to be strongly compact in itself</u> iff any sequence $(u_n,t_n)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$, $n = 1, 2, 3, \ldots$ has a subsequence (u_{n_k},t_{n_k}), $k = 1, 2, 3, \ldots$ such that $t_{n_k}\to t_o$, and $\tilde{u}_{n_k}\to u_o$ (st) as $k \to \infty$ for some $(u_o,t_o) \varepsilon \hat{\mathcal{U}}(\Omega,\Gamma)$, where $\tilde{u}_{n_k} \varepsilon L_2^m([0,t_o])$ is defined by

$$\tilde{u}_{n_k} = \begin{cases} u_{n_k} | [0,t_o], & \text{if } t_o \leq t_{n_k} \\ \bar{u}_{n_k}, & \text{if } t_{n_k} < t_o \end{cases},$$

$k = 1, 2, 3, \ldots,$ where $u_{n_k} | [0,t_o]$ denotes the restriction of u_{n_k} to $[0,t_o]$, and $\bar{u}_{n_k}$ is defined by

$$\bar{u}_{n_k}(t) = \begin{cases} u_{n_k}(t), & \text{if } 0 \leq t \leq t_{n_k} \\ u^*, & \text{if } t_{n_k} < t \leq t_o, \end{cases}$$

u^* is a fixed element of Ω .

THEOREM 4.2.1. <u>Let hypotheses</u> (4.1.3) <u>through</u> (4.1.6) <u>and the following hypotheses be fulfilled</u>:

(A) Ω <u>is a compact (nonempty) subset of</u> R^m;

(B) $\hat{\mathcal{U}}(\Omega,\Gamma) \neq \emptyset$;

(C) $\mathcal{J}:[0,T] \to \mathcal{C}(R^n)$ <u>is usc and</u> $\mathcal{J}(t) \neq \emptyset$, $0 \leq t \leq T$;

(D) $\hat{\mathcal{U}}(\Omega,\Gamma)$ <u>is strongly compact in itself</u>.

<u>Then there exists a</u> $(u_o,t_o) \varepsilon \hat{\mathcal{U}}(\Omega,\Gamma)$ <u>satisfying</u> (4.1.8).

Proof: First it will be shown that the responses $x(\cdot,u)$ to controls $(u,t_1)\in\mathcal{U}(\Omega,\Gamma)$ are uniformly bounded. To this end suppose $(u,t_1)\in\mathcal{U}(\Omega,\Gamma)$, then we have that

$$(4.2.1)\qquad \frac{d(||x(t,u)||)}{dt} \leqslant \left|\left|\frac{dx(t,u)}{dt}\right|\right|, \text{ a.e. on } [0,t_1].$$

Both $||x(\cdot,u)||$, $x(\cdot,u)$ are absolutely continuous on $[0,t_1]$. Therefore there is a set $E_o\subset[0,t_1]$ such that, the measure of E_o is zero, and $x,||x||$ are finitely differentiable on $[0,t_1]-E_o$. Now pick an interior point $t_o\in[0,t_1]$ such that $\frac{dx(t_o,u)}{dt}$ and $\frac{d(||x(t_o,u)||)}{dt}$ both exist. Consider the two difference quotients

$$\frac{||x(t,u)|| - ||x(t_o,u)||}{t-t_o},\ \frac{||x(t,u)-x(t_o,u)||}{t-t_o}$$

with $t > t_o$ sufficiently small so that $t\in[0,t_1]$. We observe that $\big|\,||x(t,u)|| - ||x(t_o,u)||\,\big| \leqslant ||x(t,u)-x(t_o,u)||$ so that

$$\frac{||x(t,u)|| - ||x(t_o,u)||}{t-t_o} \leqslant \frac{||x(t,u)-x(t_o,u)||}{t-t_o}$$

and since $x(\cdot,u),||x(\cdot,u)||$ are both finitely differentiable at t_o, we have the validity of (4.2.1) at any $t_o\in\{\text{Interior}[0,t_1]\}-E_o$, and consequently (4.2.1) is true a.e. on $[0,t_1]$. Hypothesis (4.1.5) and inequality (4.2.1) imply that

$$(4.2.2)\qquad ||x(t,u)|| \leqslant ||x_o|| + \int_0^t \mu(s)[C+||x(s,u)||]ds$$

for $t \varepsilon [0,t_1]$. From inequality (4.2.2) and a slight extension of Gronwall's Lemma [XVII, pg. 48, Exercise 1] it follows that

$$(4.2.3) \quad ||x(t,u)|| \leq [||x_o||+C] \exp \left(\int_0^t \mu(s)ds \right), \quad 0 \leq t \leq t_1 .$$

Since (4.2.3) holds for all $(u,t_1)\varepsilon\mathcal{U}(\Omega,\Gamma)$, it follows that $\{x(t,u)|0 \leq t \leq t_1, (u,t_1)\varepsilon\mathcal{U}(\Omega,\Gamma)\}$ is bounded.

The same assumptions were made on the function f^o appearing in (4.1.7) as were made on the functions f^i, $i = 1, 2, \ldots, n$ (in particular (4.1.5) holds with f replaced by $\hat{f} = (f^o,f^1,\ldots,f^n)$, so that if x is replaced by $\hat{x} = (x^o,x^1,\ldots,x^n)$, inequality (4.2.3) retains its validity), so it is immediately inferred that $\{K(u,t_1)|(u,t_1)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)\}$ is a nonempty bounded set of real numbers. Consequently

$$(4.2.4) \qquad \inf \{K(u,t_1)|(u,t_1)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)\} \underset{d}{\equiv} \gamma, \quad +\infty < \gamma < -\infty .$$

Select a "minimizing sequence: $(u_n,t_n)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$, $n = 1, 2, 3, \ldots$, i.e., let $(u_n,t_n)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$, $n = 1, 2, 3, \ldots$ be such that

$$(4.2.5) \qquad K(u_n,t_n) \rightarrow \gamma \quad \text{as} \quad n \rightarrow \infty .$$

There is a subsequence of $\{(u_n,t_n)\}$, which we shall still call $\{(u_n,t_n)\}$, such that $t_n \rightarrow$ some $t_o\varepsilon\Gamma$ (monotonely) as $n \rightarrow \infty$. There are two cases:

<u>Case 1</u> $\qquad 0 \leq t_o \leq t_n \leq T, \quad n = 1, 2, 3, \ldots$;

<u>Case 2</u> $\qquad 0 \leq t_n \leq t_o \leq T, \quad n = 1, 2, 3, \ldots$.

We shall dispose of Case I first. Since $\hat{\mathcal{U}}(\Omega,\Gamma)$ is strongly compact in itself, we may select a further subsequence of $\{(u_n,t_n)\}$ (still denoted by $\{(u_n,t_n)\}$) such that $u_n|[0,t_o]$ (the restriction of u_n to $[0,t_o]$) converges strongly to some u_o such that $(u_o,t_o)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$. It is therefore follows that (see [XI, pg. 87]) there exists a further subsequence of $\{(u_n,t_n)\}$ (still called $\{(u_n,t_n)\}$) such that

(4.2.6) $u_n(t)\rightarrow u_o(t)$ as $n\rightarrow\infty$ a.e. on $[0,t_o]$, and $t_n\rightarrow t_o$ (monotonely) as $n\rightarrow\infty$.

The responses corresponding to the sequence of controls in (4.2.6) are given by

$$x(t,u_n) = x_o + \int_0^t f(\xi,x(\xi,u_n),u_n(\xi))d\xi, \tag{4.2.7}$$

$0\leqslant t\leqslant t_n$, $n = 1, 2, 3, \ldots$. It is our intention to show that the family of functions, $\{x(\cdot,u_n),\ n = 1, 2, 3, \ldots\}$ is equicontinuous on the interval $[0,t_o]$. Thus let $t_1,t_2\varepsilon[0,t_o]$ and observe that

$$||x(t_1,u_n)-x(t_2,u_n)||\leqslant \int_{t_m}^{t_M} ||f(\xi,x(\xi,u_n),u_n(\xi))||d\xi ,$$

where $t_M = \max(t_1,t_2)$ and $t_m = \min(t_1,t_2)$. Consequently by hypothesis (4.1.5)

$$||x(t_1,u_n)-x(t_2,u_n)||\leqslant \int_{t_m}^{t_M} \mu(\xi)[C+||x(\xi,u_n)||]d\xi .$$

From (4.2.3) it follows that there is a constant $H>0$ such that $||x(t,u_n)||<H$, $n = 1, 2, 3, \ldots$, $0\leq t\leq t_o$. Therefore we obtain that

$$(4.2.8) \qquad ||x(t_1,u_n)-x(t_2,u_n)||\leq[H+C] \int_{t_m}^{t_M} \mu(\xi)d\xi \;.$$

But the function ϕ defined by $\phi(t) = \int_0^t \mu(\xi)d\xi$, $0\leq t\leq t_o$, is absolutely continuous on $[0,t_o]$, a fortiori uniformly continuous on $[0,t_o]$, so that given $\varepsilon>0$ there is a $\delta>0$ such that

$$|t_1-t_2|<\delta,\ t_1,\ t_2\varepsilon[0,t_o] \Rightarrow ||x(t_1,u_n)-x(t_2,u_n)||<\varepsilon,$$

for $n = 1, 2, 3, \ldots$. Whence $\{x(\cdot,u_n)\}$ is a uniformly bounded equicontinuous family of functions on $[0,t_o]$. Thus by the Ascoli-Arzelà Theorem [X, pg. 61] there is a subsequence of the $\{x(\cdot,u_n)\}$, and a corresponding subsequence of the sequence $\{(u_n,t_n)\}$ (which we still call $\{x(\cdot,u_n)\}$ and $\{(u_n,t_n)\}$, respectively) such that

$x(t,u_n)\rightarrow\bar{x}(t)$ uniformly on $[0,t_o]$ as $n\rightarrow\infty$, and $u_n\rightarrow u_o$ (a.e.), $t_n\rightarrow t_o$ as $n\rightarrow\infty$. Now

$$x(t,u_n) = x_o + \int_0^t f(\xi,x(\xi,u_n),u_n(\xi))d\xi \;,$$

$0\leq t\leq t_o$, and $||f(\xi,x(\xi,u_n),u_n(\xi))||\leq[H+C]\mu(\xi)$, $0\leq\xi\leq t_o$, so by the Lebesgue dominated convergence theorem we obtain that

$$(4.2.9) \qquad \bar{x}(t) = x_o + \int_0^t f(\xi,\bar{x}(\xi),u_o(\xi))d\xi,$$

$0 \leq t \leq t_o$, and therefore $\bar{x}$ is the unique response to the control u_o having domain $[0, t_o]$. It follows that

$$\bar{x}(t) = x(t, u_o), \quad 0 \leq t \leq t_o.$$

It must now be shown that

$$x(t_n, u_n) \to x(t_o, u_o) \quad \text{as} \quad n \to \infty . \tag{4.2.10}$$

In order to accomplish this we first observe that

$$||x(t_o, u_n) - x(t_n, u_n)|| \leq \int_{t_o}^{t_n} ||f(\xi, x(\xi, u_n), u_n(\xi))|| d\xi$$

$$\leq \int_{t_o}^{t_n} [H+C] \mu(\xi) d\xi ,$$

and since $t_n \to t_o$ as $n \to \infty$, the last inequality gives

$$\lim_{n \to \infty} x(t_n, u_n) = \lim_{n \to \infty} x(t_o, u_n) . \tag{4.2.11}$$

It follows from relation (4.2.11) that

$$\lim_{n \to \infty} [x(t_n, u_n) - x(t_n, u_n) + x(t_o, u_n)] = \lim_{n \to \infty} x(t_n, u_n) = \lim_{n \to \infty} x(t_o, u_n) = x(t_o, u_o)$$

(note that it has already been demonstrated that $x(t, u_n) \to x(t, u_o)$ as $n \to \infty$, $0 \leq t \leq t_o$). Consequently the validity of relation (4.2.10) has been established. Thus we have $t_n \to t_o$ as $n \to \infty$, $x(t_n, u_n) \in \mathfrak{J}(t_n)$, $n = 1, 2, 3, \ldots,$ and relation (4.2.10), so from Theorem 2.1.3 it follows that

$$x(t_o, u_o) \in \mathfrak{J}(t_o),$$

thereby showing that $(u_o,t_o)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$.

Finally we have that

$$K(u_n,t_n) = \int_0^{t_n} f^o(\xi,x(\xi,u_n),u_n(\xi))d\xi \to K(u_o,t_o) = \gamma$$

as $n\to\infty$ (by (4.2.6), (4.2.9), the continuity of f^o, and the Lebesgue dominated convergence theorem). This proves the theorem for Case 1. The details of the proof for Case 2 are similar enough to those exhibited for Case 1, so that we shall not supply the details of the proof of the theorem for Case 2.

In the remainder of this chapter we shall change the definition of $\mathcal{U}(\Omega,\Gamma)$ several times. The sense in which it is to be understood will be announced in each of the remaining corollaries.

COROLLARY 4.2.1. Suppose hypotheses (4.1.3) through (4.1.6) and the following hypotheses are satisfied:

(A) Ω is a compact (nonempty) subset of R^m ;

(B) $\hat{\mathcal{U}}(\Omega,\Gamma) \neq \emptyset$;

(C) The mapping $\mathcal{J}:[0,T]\to\mathcal{B}(R^n)$ is usc and $\mathcal{J}(t) \neq \emptyset$, $0\leq t\leq T$;

(D) There is a constant $\Delta>0$ such that $(u,t_1)\varepsilon\mathcal{U}(\Omega,\Gamma)$ implies the total variation of u^i over $[0,t_1]$ is less than or equal to Δ, $i = 1, 2, \ldots, m$;

(E) $\{\bar{u}|(u,t_1)\varepsilon\mathcal{U}(\Omega,\Gamma)\}$ is a closed subset of $L_2^m([0,T])$ (where for $(u,t_1)\varepsilon\mathcal{U}(\Omega,\Gamma)$, $\bar{u}\equiv u$ on $[0,t_1]$, $\bar{u}\equiv 0$ on $R-[0,t_o]$). Then there is a $(u_o,t_o)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$ such that (4.1.8) is satisfied.

Proof: $\hat{\mathcal{U}}(\Omega,\Gamma)$ is strongly compact in itself. For if $(u_n,t_n)\in\hat{\mathcal{U}}(\Omega,\Gamma)$, $n = 1, 2, 3, \ldots$, then there is a subsequence of $\{(u_n,t_n)\}$ (call it $\{(u_{n_k},t_{n_k})\}$) such that $t_{n_k}\to t_o$ as $k\to\infty$, for some $t_o\in\Gamma$ (the convergence may be assumed to be monotone). Then by Helly's compactness theorem [XII, pg. 119], there is a subsequence of $\{(\tilde{u}_{n_k},t_{n_k})\}$ (see Definition 4.2.1 for the meaning of $\tilde{u}_{n_k}$; in extracting the subsequence of $\{(\tilde{u}_{n_k},t_{n_k})\}$ we retain the same notation) such that

$$\tilde{u}_{n_k}(t)\to u_o(t),\quad 0<t<t_o,\quad t_{n_k}\to t_o\in\Gamma \quad \text{as} \quad k\to\infty.$$

Moreover, the function u_o must satisfy hypothesis (D) of this theorem, and obviously $u_o(t)\in\Omega$, for each $t\in[0,t_o]$, since Ω is compact. Whence by the Lebesgue dominated convergence theorem

$$\int_0^{t_o} [u_{n_k}(\xi)-u_o(\xi)]^2 d\xi\to 0 \quad \text{as} \quad k\to\infty.$$

Thus by hypothesis (E) of this theorem $(u_o,t_o)\in\hat{\mathcal{U}}(\Omega,\Gamma)$. Therefore $\hat{\mathcal{U}}(\Omega,\Gamma)$ is strongly compact in itself, and the desired conclusion follows from Theorem 4.2.1.

COROLLARY 4.2.2. Suppose hypotheses (4.1.3) through (4.1.6) and the following hypotheses are satisfied:

(A) Ω is a compact (nonempty) subset of R^m;

(B) $\hat{\mathcal{U}}(\Omega,\Gamma) \neq \emptyset$;

(C) For every $\varepsilon>0$, there exists $\delta>0$ such that for all $(u,t_1)\in\hat{\mathcal{U}}(\Omega,\Gamma)$ and $|h|<\delta$

$$\int_0^T ||\overline{u}(\xi+h)-\overline{u}(\xi)||^2 d\xi < \varepsilon$$

(*see hypothesis* (E) *of Corollary* 4.2.1 *for the meaning of* $\overline{u}$);

(D) *The mapping* $\mathcal{J}:[0,T]\rightarrow (R^n)$ *is usc and* $\mathcal{J}(t)\neq\emptyset$, $0\leq t\leq T$;

(E) $\{\overline{u}\ (u,t_1)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)\}$ *is a closed subset of* $L_2^m([0,T])$;

Then there is a $(u_o,t_o)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$ *satisfying* (4.8.1).

Proof: $\hat{\mathcal{U}}(\Omega,\Gamma)$ is strongly compact in itself [XII, pg. 44]. The conclusion follows from Theorem 4.2.1.

COROLLARY 4.2.3. *Retain all the hypotheses in Corollary* 4.2.2 *with the exception of* (C), *and for hypothesis* (C) *substitute*

(C') *There is a constant* $A>0$ *such that* $(u,t_1)\varepsilon\mathcal{U}(\Omega,\Gamma)$ *implies*

$$||u(t)-u(t')||\leq A|t-t'|,\quad 0\leq t,\ t'\leq t_1$$

(*uniformly with respect to* $(u,t_1)\varepsilon\mathcal{U}(\Omega,\Gamma)$).

Then there is a $(u_o,t_o)\varepsilon\hat{\mathcal{U}}(\Omega,\Gamma)$ *satisfying* (4.1.8).

Proof: Hypothesis (C') of this corollary implies hypothesis (C) of Corollary 4.2.2, so that the conclusion follows from Corollary 4.2.2.

CHAPTER V

THE EXISTENCE OF OPTIMAL CONTROLS FOR SYSTEMS WITH A DELAYED ARGUMENT

In this chapter we extend a theorem of E. Roxin [17] to include the optimal control problem in a system with a delayed argument.

5.1 Formulation of the Optimal Control Problem in a System with a Delayed Argument.

The control system is given by a system of real ordinary differential-difference equations of the following type

$$\dot{x}(t) = f(t,x(t),x(t-\omega), u(t))), \tag{5.1.1}$$

where $\omega>0, x\in R^n$, and $u\in R^m$. The following assumptions are made:

(A) The mapping $f:R\times R^n\times R^m \rightarrow R^n$ is continuous;

(B) Ω is a fixed compact subset of R^m;

(C) A Lipschitz condition is satisfied. There is a constant $L>0$ such that if (t,x,y,u), $(t,x_1,y_1,u)\in R\times R^n\times R^n\times R^m$ we have that

$$||f(t,x,y,u)-f(t,x_1,y_1,u)||\leq L[||x-x_1||+||y-y_1||] ;$$

(D) For all $u\epsilon\Omega$ (uniformly),

$$||f(t,x,y,u)||\leq\mu(t)[D+||x||+||y||],$$

where the function μ is integrable in every finite interval, and D is a positive constant;

(E) We define the set $C = \{\phi|\phi:[0,\omega]\to R^n,\ \phi$ continuous$\}$ with norm defined by $||\phi|| = \sup\{||\phi(t)||\ |0\leq t\leq\omega\}$, when $\phi\epsilon C$. Let H be some compact subset of the normed linear space C.

(F) Given an initial function $\phi\epsilon H\subset C$ we require that

$$x(t) = \phi(t), \quad 0\leq t\leq\omega .$$

(G) A set $\mathcal{U}(\Omega)$ will be defined by saying $u\epsilon\mathcal{U}(\Omega)$ iff u has domain $[\omega,t_1]$ for some $[\omega,t_1]\subset[0,T]$, where $T>0$ is constant, u is measurable on $[\omega,t_1]$, and $u(t)\epsilon\Omega$ for each $t\epsilon[\omega,t_1]$;

(H) $f(t,x,y,\Omega) = \{f(t,x,y,u)|u\epsilon\Omega\}$ is a convex subset of R^n. We note that since f is continuous in u for each $(t,x,y)\epsilon R\times R^n\times R^n$ and Ω is compact, it follows that $f(t,x,y,\Omega)$ is also compact for each $(t,x,y)\epsilon R\times R^n\times R^n$.

The above conditions are sufficient to guarantee that corresponding to any $u\epsilon\mathcal{U}(\Omega)$, u having domain $[\omega,t_1]\subset[0,T]$, and any initial function $\phi\epsilon H$, there exists a unique absolutely continuous function $x(\cdot,\phi,u)$ mapping $[\omega,t_1]\to R^n$ satisfying (5.1.1) a.e. on $[\omega,t_1]$ together with the initial (functional) condition,

$$(5.1.2) \qquad x(t,\phi,u) = \phi(t), \quad 0 \le t \le \omega .$$

The proof of this last assertion is obtained in the usual way [III, Chapter 6] by continuing the solution from interval to interval. A solution on $[\omega, \min(t_1,2\omega)]$ exists by standard existence theorems [XIII, pg. 342], and uniqueness on $[\omega, \min(t_1,2\omega)]$ follows as in [XIII, pp. 345-346].

DEFINITION 5.1.1. A point $x_1 \varepsilon R^n$ is said to be <u>attainable by a pair of functions</u> $(\phi,u)\varepsilon H\times\mathcal{U}(\Omega)$ iff the unique solution $x(\cdot,\phi,u):[\omega,t_1]\to R^n$ (where $[\omega,t_1]$ is the domain of u) of (5.1.1) corresponding to $(\phi,u)\varepsilon H\times\mathcal{U}(\Omega)$ satisfies

$$x(t_1,\phi,u) = x_1.$$

The attainable set $\mathcal{Q}$ is defined to be

$$\mathcal{Q} = \{x\varepsilon R^n \mid x \ \underline{\text{is attainable by some}}\ (\phi,u)\varepsilon H\times\mathcal{U}(\Omega)\} .$$

For $t\varepsilon[\omega,T]$ we define the fixed-time cross-section of $\mathcal{Q}$ at time t to be

$$\mathcal{Q}_t \equiv \{x\varepsilon R^n \mid x \ \underline{\text{is attainable by}}\ (\phi,u)\varepsilon H\times(\mathcal{U}(\Omega)\,|\,[\omega,t])\},$$

where $\mathcal{U}(\Omega)\,|\,[\omega,t] \equiv \{u \mid u\varepsilon\mathcal{U}(\Omega)$ such that u has domain $[\omega,t]\}$. Let $F:[\omega,T]\to \mathcal{C}(R^n)$ be usc (see Chapter II for the meaning of the terminology). Define

$$\hat{\mathcal{U}}(t_1,\Omega,H) = \{(\phi,u)\,|\,(\phi,u)\varepsilon H\times\mathcal{U}(\Omega),\ u \ \underline{\text{has domain}}\ [\omega,t_1]\ \underline{\text{and}}\ x(t_1,\phi,u)\varepsilon F(t_1)\}\ .$$

for $\omega\leq t_1\leq T$, and then we define

$$\hat{\mathcal{U}}(H,\Omega) = \bigcup_{\omega\leq t\leq T}\hat{\mathcal{U}}(t,H,\Omega).$$

The problem which we shall consider in this chapter is

$$(5.1.3)\qquad x^1(t_1,\phi,u) = \min \text{ (or max) on } \hat{\mathcal{U}}(H,\Omega),$$

where $x^1(t_1,\phi,u)$ denotes the first component of $x(\cdot,\phi,u)$ evaluated at the terminal time t_1 of the control function u.

5.2 An Existence Theorem for the Optimal Control Problem in a System with a Delayed Argument.

THEOREM 5.2.1. *If the system* (5.1.1) *satisfies hypotheses* (A)-(H), *then the attainable set* $\mathcal{R}$ *is closed.*

We shall first prove a lemma.

LEMMA 5.2.1. *If the system* 5.1.1 *satisfies hypotheses* (A)-(H), *then* $\mathcal{R}_{t_1}$ *is closed for each* $t_1\varepsilon[\omega,T]$.

Proof of Lemma 5.2.1: If $\mathcal{R}_{t_1} = \emptyset$, the conclusion is obvious. Thus assume $\mathcal{R}_{t_1} \neq \emptyset$, and let $\xi_1,\xi_2,\xi_3,\ldots$ be a sequence of points in $\mathcal{R}_{t_1}$ such that $\xi_i\to\xi_o$ as $i\to\infty$. We shall prove $\xi_o\varepsilon\mathcal{R}_{t_1}$. Since $\xi_i\varepsilon\mathcal{R}_{t_1}$, $i = 1, 2, 3, \ldots$, there is a sequence

$(\phi_i,u_i)\in H\times\mathcal{U}(\Omega)$, u_i having domain $[\omega,t_1]$, for $i = 1, 2, 3, \ldots$ such that

(5.2.4) $x(t,\phi_i,u_i) =$

$$\phi_i(\omega)+\int_\omega^t f(s,x(s,\phi_i,u_i),x(s-\omega,\phi_i,u_i),u_i(s))ds \ ,$$

for $\omega\leq t\leq t_1, x(t_1,\phi_i,u_i) = \xi_i$, and $x(t,\phi_i,u_i) = \phi_i(t)$ whenever $0\leq t\leq\omega$, $i = 1, 2, 3, \ldots$.

We shall prove $x(t,\phi_i,u_i)$, $i = 1, 2, 3, \ldots$ is a bounded sequence (uniformly with respect to $t\in[\omega,t_1]$). It is easy to establish that

(5.2.5) $$\frac{d||x(t,\phi_i,u_i)||}{dt}\leq\left|\left|\frac{dx(t,\phi_i,u_i)}{dt}\right|\right|, \text{ a.e. on } [\omega,t_1].$$

Define $H[t] = \{\phi(t)|\phi\in H\}$, for $0\leq t\leq\omega$. Then since H is a compact subset of C, it can be shown that $S = \bigcup_{0\leq t\leq\omega}H(t)$ is a bounded subset of R^n. Using hypothesis (D) and (5.2.5) we have that

(5.2.6) $$\frac{d||x(t,\phi_i,u_i)||}{dt} \leq$$

$$\mu(t)[D+||x(t,\phi_i,u_i)||+||x(t-\omega,\phi_i,u_i)||)$$

a.e. on $[\omega,t_1]$, $i = 1, 2, 3, \ldots$.

In view of hypothesis (F) (5.2.6) becomes

$$\frac{d||x(t,\phi_i,u_i)||}{dt} \leq \mu(t)[D+||x(t,\phi_i,u_i)||+||\phi_i(t)||], \tag{5.2.7}$$

for almost every $t\varepsilon[\omega, \min(2\omega,t_1)]$, $i = 1, 2, 3, \ldots$.

But since the set $S = \bigcup_{0\leq t\leq\omega} H[t]$ is bounded, there is constant $B>0$ such that $||\phi(t)||\leq B$, for each $t\varepsilon[0,\omega]$, and for each $\phi\varepsilon H$. Therefore

$$\frac{d||x(t,\phi_i,u_i)||}{dt} \leq \mu(t)[D+B+||x(t,\phi_i,u_i)||], \tag{5.2.8}$$

a.e. on $[\omega, \min(2\omega, t_1)]$, $i = 1, 2, 3, \ldots$. Whence

$$\int_\omega^t \frac{d}{ds}(\ln(D+B+||x(s,\phi_i,u_i)||)ds \leq \int_\omega^t \mu(s)\,ds,$$

for $\omega\leq t\leq \min(2\omega,t_1)$, $i = 1, 2, 3, \ldots$, from which we obtain that $\ln(D+B+||x(t,\phi_i,u_i)||)\leq h(t)$, $\omega\leq t\leq\min(2\omega,t_1)$, $i = 1, 2, 3, \ldots$, where

$$h(t) = \ln(D+B+||\phi_i(\omega)||) + \int_\omega^t \mu(s)\,ds,$$

and we have that $0\leq h(t)\leq \ln(2B+1)+\int_\omega^T \mu(s)\,ds = N = \text{constant}$,

for each t in $[\omega, \min(2\omega,t_1)]$. Consequently for each

$t\varepsilon[\omega, \min(2\omega,t_1)]$ we have that $\ln(D+B+||x(t,\phi_i,u_i)||)\leq N$, $i = 1, 2, 3, \ldots,$ and it thereby follows that

$$(5.2.9) \qquad ||x(t,\phi_i,u_i)||\leq(e^N-D-B) \ ,$$

for $\omega\leq t\leq\min(2\omega,t_1)$, $i = 1, 2, 3, \ldots$. Define $I_k =$ $\{t|\omega\leq t\leq \min(k\omega,t_1)\}$, $k = 2, 3, \ldots,$ then there is a least integer L such that $\bigcup_1^L I_k \supset [\omega,t_1]$. We shall show by induction that there are constants $A_1,A_2,\ldots,A_L$ such that

$$||x(t,\phi_i,u_i)||\leq A_k, \ \omega\leq t\leq\min(k\omega,t_1), \ i = 1, 2, 3, \ldots .$$

Assume that k such constants have been found, we shall show how to construct A_{k+1} (note that by (5.2.9), A_1 can be taken to be (e^N-D-B)). By the induction hypothesis we have that $||x(t,\phi_i,u_i)||\leq A_k$, $\omega\leq t\leq \min(k\omega,t_1)$, $i = 1, 2, 3, \ldots$. Thus from inequality (5.2.6) we ascertain that

$$(5.2.10) \qquad \frac{d||x(t,\phi_i,u_i)||}{dt}\leq\mu(t)[D+A_k+||x(t,\phi_i,u_i)||]$$

a.e. on $[\omega, \min(k+1)\omega,t_1)]$, $i = 1, 2, 3, \ldots$. Consequently by reasoning similar to that used in proving inequality (5.2.9) we are able to show that

$$\ln(D+A_k+||x(t,\phi_i,u_i)||)\leq\ln(2A_k+D)+\int_\omega^T \mu(s)\,ds.$$

Thus if we define $N_{k+1} = \ln(D+2A_k)+\int_\omega^T \mu(s)\,ds$ = constant, then the constant A_{k+1} defined by

$$A_{k+1} = (e^{N_{k+1}} - D - A_k)$$

has the desired properties. Therefore it follows by induction that

$$(5.2.11) \qquad ||x(t,\phi_i,u_i)|| \le A_L,\ \omega \le t \le t_1,\ i = 1, 2, 3, \ldots,$$

where L is the positive integer defined above.

We define a sequence of functions as follows:

$$(5.2.12) \qquad r_i(t) = f(t,x(t,\phi_i,u_i),x(t-\omega,\phi_i,u_i),u_i(t))$$

for $\omega \le t \le t_1$, $i = 1, 2, 3, \ldots$. Then since $u_i(t) \varepsilon \Omega$ for $\omega \le t \le t_1$, Ω is compact, and f is continuous on $R \times R^n \times R^n \times R^m$, it follows from (5.2.11) that the sequence $\{r_i(t)\}$ is bounded (uniformly with respect to $t \varepsilon [\omega,t_1]$). Thus $r_i \varepsilon L_2^n([\omega,t_1])$, $i = 1, 2, 3, \ldots$, and the sequence of L_2^n-norms, $\{||r_i||\}$, is bounded. Consequently [XII, pg. 117], there is a $r_o \varepsilon L_2^n([\omega,t_1])$ and a subsequence of $\{r_i\}$ (still denoted by $\{r_i\}$) such that

$$(5.2.13) \qquad r_i \to r_o \text{ (wk) as } i \to \infty.$$

We then select a subsequence of $\{\phi_i\}$, and the corresponding subsequence of $\{r_i\}$ (still without changing the notation) such that $\phi_i \to$ some $\phi_o \varepsilon H$ uniformly on $[0,\omega]$ as $i \to \infty$. Define

$$x_o(t) = \begin{cases} \phi_o(\omega) + \int_\omega^t r_o(s)\,ds, & \omega \leq t \leq t_1 \\ \phi_o(\omega), & 0 \leq t \leq \omega \end{cases} \quad . \tag{5.2.14}$$

In view of (5.2.13), (5.2.14), and the fact that $\phi_i \to \phi_o$ uniformly on $[0,\omega]$ as $i \to \infty$ we have that $x(t,\phi_i,u_i) \to x_o(t)$ as $i \to \infty$, $0 \leq t \leq t_1$, a fortiori

$$x(t_1,\phi_i,u_i) = \xi_i \to \xi_o = x_o(t_1). \tag{5.2.15}$$

We shall show that $r_o(t) \varepsilon f(t,x_o(t),x_o(t-\omega),\Omega)$ for almost every $t \varepsilon [\omega,t_1]$. As a first step in demonstrating the validity of this assertion we show that for any $a \varepsilon R^n$

$$\limsup_{i \to \infty} \langle a,r_i(t)\rangle \geq \langle a,r_o(t)\rangle \geq \liminf_{i \to \infty} \langle a,r_i(t)\rangle \quad , \tag{5.2.16}$$

almost everywhere on $[\omega,t_1]$. Here as elsewhere in this dissertation, $\langle x,y\rangle$ denotes the "inner product" of x and y, x, $y \varepsilon R^n$. Suppose that

$$\limsup_{i \to \infty} \langle a,r_i(t)\rangle < \langle a,r_o(t)\rangle$$

on some set $E \subset [\omega,t_1]$ of positive measure. Then from the fact that $||r_i(t)|| \leq (2A_L+D)\mu(t)$, a.e. on $[\omega,t_1]$ (by (5.2.6), (5.2.11) and hypothesis (D)) it follows from Fatou's lemma [XIII, pg. 167] that

$$\limsup_{i\to\infty}\left\langle a, \int_E r_i(s)\,ds\right\rangle \leq \int_E \limsup_{i\to\infty}\langle a, r_i(s)\rangle\,ds < \left\langle a, \int_E r_o(s)\,ds\right\rangle .$$

But by (5.2.13)

$$\limsup_{i\to\infty}\left\langle a, \int_E r_i(s)\,ds\right\rangle = \lim_{i\to\infty}\left\langle a, \int_E r_i(s)\,ds\right\rangle = \left\langle a, \int_E r_o(s)\,ds\right\rangle .$$

From this contradiction we infer that the first half of inequality (5.2.16) is valid. The other half of inequality (5.2.16) is proved dually. Since f is continuous in the variable $u\varepsilon\Omega$ and Ω is compact, it follows that $f(t,x(t,\phi_i,u_i),x(t-\omega,\phi_i,u_i),\Omega)$ is compact, $i = 1, 2, 3, \ldots$. Therefore there exist $v_i,\overline{v}_i\varepsilon\Omega$ such that

$$\sup\{\langle a,f(t,x(t,\phi_i,u_i),x(t-\omega,\phi_i,u_i),\Omega)\rangle\} =$$

$$\langle a,f(t,x(t,\phi_i,u_i),x(t-\omega),\phi_i,u_i),\overline{v}_i)\rangle ,$$

and

$$\inf\{\langle a,f(t,x(t,\phi_i,u_i)x(t-\omega,\phi_i,u_i),\Omega)\rangle\} =$$

$$\langle a,f(t,x(t,\phi_i,u_i),x(t-\omega,\phi_i,u_i),v_i)\rangle ,$$

$i = 1, 2, 3, \ldots$. Since $u_i(t)\varepsilon\Omega$, $\omega\leq t\leq t_1$, $i = 1, 2, 3, \ldots$

we have that

$$\langle a,f(t,\phi_i,u_i),x(t-\omega,\phi_i,u_i),\overline{v}_i)\rangle \geq \langle a,r_i(t)\rangle \geq$$

$$\langle a,f(t,x(t,\phi_i,u_i),x(t-\omega,\phi_i,u_i),v_i)\rangle ,$$

for $t\varepsilon[\omega,t_1]$. But Ω is a compact subset of R^m, so that without loss of generality we may suppose that we actually have

$$v_i \to v_o, \overline{v}_i \to \overline{v}_o, \quad \text{for some } v_o, \overline{v}_o \varepsilon \Omega \text{ as } i \to \infty .$$

Thus for $t\varepsilon[\omega,t_1]$ such that (5.2.16) is valid, we obtain that

$$\limsup_{i\to\infty} \langle a, f(t,x(t,\phi_i,u_i),x(t-\omega,\phi_i,u_i),\overline{v}_i)\rangle \geq$$

$$\langle a, r_o(t)\rangle \geq \liminf_{i\to\infty} \langle a, f(t,x(t,\phi_i,u_i),x(t-\omega,\phi_i,u_i),v_i)\rangle .$$

Therefore by the continuity of f we have

$$(5.2.17) \quad \langle a, f(t,x_o(t),x_o(t-\omega),\overline{v}_o)\rangle \geq \langle a, r_o(t)\rangle \geq \langle a, f(t,x_o(t),x_o(t-\omega),v_o)\rangle .$$

Thus for almost every $t\varepsilon[\omega,t_1]$ we have that

$$(5.2.18) \qquad \sup \langle a, f(t,x_o(t),x_o(t-\omega),\Omega)\rangle \geq$$

$$\langle a, r_o(t)\rangle \geq \inf \langle a, f(t,x_o(t),x_o(t-\omega),\Omega)\rangle .$$

The inequalities in (5.2.18) are valid for each $a\varepsilon R^n$. Thus let $a\varepsilon R^n$, $b\varepsilon R$ be such that

$$P = \{x \mid \langle a,x\rangle \geq b, x\varepsilon R^n\} \supset f(t,x_o(t),x_o(t-\omega),\Omega),$$

then by (5.2.17) we have $r_o(t)\varepsilon P$. Also if

$$P = \{x \mid \langle a,x\rangle \leq b, x\varepsilon R^n\} \supset f(t,x_o(t),x_o(t-\omega),\Omega)$$

we once again have $r_o(t)\varepsilon P$. Whence we have shown that every closed halfspace containing the compact convex set $f(t,x_o(t),x_o(t-\omega),\Omega)$ also contains $r_o(t)$ (this assertion is true for almost every $t\varepsilon[\omega,t_1]$). This, however, is sufficient to prove that

$$r_o(t)\varepsilon f(t,x_o(t),x_o(t-\omega),\Omega), \quad \text{a.e. on} \quad [\omega,t_1] ,$$

(consult [VII, pg. 23]). Whence for almost every $t\varepsilon[\omega,t_1]$ there is a $u_o\varepsilon\Omega$ such that

$$r_o(t) = f(t,x_o(t),x_o(t-\omega),u_o) .$$

In this way a function u_o can be defined at almost every point of the interval $[\omega,t_1]$. Obviously we can extend the definition to the remaining points of the interval in any convenient way such that $u_o(t)\varepsilon\Omega$. Thus u_o has its range in Ω, and by Filippov's lemma [7, pp. 78-79], there is a measurable function $u_o:[\omega,t_1]\rightarrow\Omega$ such that

$$(5.2.19) \quad r_o(t) = f(t,x_o(t),x_o(t-\omega),u_o(t)), \quad \text{a.e. on} \quad [\omega,t_1].$$

Whence

$$x_o(t) = \begin{cases} \phi_o(\omega)+\displaystyle\int_\omega^t f(s,x_o(s),x_o(s-\omega),u_o(s))\,ds, & \omega\leq t\leq t_1 \\ \phi_o(t), & 0\leq t\leq\omega \end{cases} ,$$

so that we may write using the notation adopted in this chapter that

$$x_0(t) = x(t,\phi_0,u_0), \quad 0\leq t\leq t_1.$$

Therefore from (5.2.16) it follows that

$$x_0(t_1) = x(t_1,\phi_0,u_0) = \xi_0 \varepsilon \mathcal{R}_{t_1} ,$$

thereby proving that $\mathcal{R}_{t_1}$ is closed.

Proof of Theorem 5.2.1: We assume $\mathcal{R} \neq \emptyset$, because we know that the empty set is closed. Moreover,

$$\mathcal{R} = \{\mathcal{R}_t | \omega\leq t\leq T\} ,$$

and we shall adopt the notational device $(t_i,\xi_i)\varepsilon\mathcal{R}$ means that $\xi_i \varepsilon \mathcal{R}_{t_1}$. Thus let $(t_i,\xi_i)\varepsilon\mathcal{R}$ and suppose that $(t_i,\xi_i)\rightarrow(t_1,\xi_1)$ as $i\rightarrow\infty$. We must show that $(t_1,\xi_1)\varepsilon\mathcal{R}$, i.e., $\xi_1\varepsilon\mathcal{R}_{t_1}$.

There is a sequence $(\phi_i,u_i)\varepsilon H\times\mathcal{U}(\Omega)$ such that u_i has domain $[\omega,t_i]$ and $x(t_i,\phi_i,u_i) = \xi_i$, $i = 1, 2, 3, \ldots$. A subsequence of $\{(t_i,\xi_i)\}$ (still called $\{(\xi_i,t_i)\}$) may be selected such that $t_i\rightarrow t_1$ (monotonely) as $i\rightarrow\infty$. There are two cases:

<u>Case 1</u> $t_i\geq t_1$, $i = 1, 2, 3, \ldots$;

<u>Case 2</u> $t_i\leq t_1$, $i = 1, 2, 3, \ldots$.

We shall assume <u>Case 1</u> obtains and leave <u>Case 2</u> for the reader. We define a sequence

$$(5.2.20) \quad \eta_i = \phi_i(\omega)+\int_{\omega}^{t_1} f(s,x(s,\phi_i,u_i),x(s-\omega,\phi_i,u_i),u_i(s))\, ds,$$

$i = 1, 2, 3, \ldots$. Then

$$||\eta_i - x(t_i,\phi_i,u_i)|| \leq \int_{t_1}^{t_i} ||f(s,\phi_i,u_i),x(s-\omega,\phi_i,u_i),u_i(s))||ds,$$

and the integrand on the left hand side of this last inequality is bounded (this assertion is justified by an argument entirely similar to that used in the proof of Lemma 5.2.1). Consequently $||\eta_i - x(t_i,\phi_i,u_i)|| = ||\eta_i - \xi_i|| \to 0$ as $i \to \infty$. Therefore $\lim_{i\to\infty} \eta_i = \lim_{i\to\infty} \xi_i = \xi_1$, and from (5.2.20) $\eta_i \epsilon \mathcal{R}_{t_1}$, $i = 1, 2, 3, \ldots$.

But $\mathcal{R}_{t_1}$ is closed by Lemma 5.2.1, thereby showing that $\xi_1 \epsilon \mathcal{R}_{t_1}$, or $(t_1,\xi_1) \epsilon \mathcal{R}$, which completes the proof of the theorem.

Define

$$Q = \{x(t_1,\phi,u) | (\phi,u) \epsilon \hat{\mathcal{U}}(H,\Omega)\},$$

then we have the theorem.

THEOREM 5.2.2. <u>If the system</u> (5.1.1) <u>satisfies conditions</u> (A)-(H), <u>and</u> $Q \neq \emptyset$, <u>then there is a</u> $(\phi,u) \epsilon \hat{\mathcal{U}}(H,\Omega)$ <u>satisfying</u> (5.1.3).

Proof: First we shall show that Q is compact. Since Q is bounded, it will suffice to show that Q is closed. Let $\xi_i \epsilon Q$, $i = 1, 2, 3, \ldots$ be a sequence such that $\xi_i \to \xi_1$ as $i \to \infty$. Then there exist $(\phi_i,u_i) \epsilon \hat{\mathcal{U}}(H,\Omega)$, where u_i has domain $[\omega,t_i] \subset [\omega,T]$, $i = 1, 2, 3, \ldots$ such that

$$\xi_i = x(t_i,\phi_i,u_i) \epsilon F(t_i), \quad i = 1, 2, 3, \ldots .$$

There is a subsequence of $\{t_i\}$ and a corresponding subsequence of $\{\xi_i\}$ (still denoted by $\{t_i\}$ and $\{\xi_i\}$ respectively) such that $t_i \to t_1$ as $i \to \infty$. Then we have the following situation

$$\xi_i \varepsilon F(t_i),\ i = 1, 2, 3, \ldots,\quad \xi_i \to \xi_1, t_i \to t_1 \text{ as } i \to \infty.$$

Thus since F is usc on $[\omega,T]$ it follows from Theorem 2.1.3 that $\xi_1 \varepsilon F(t_1)$. Now $\mathcal{R}$ is closed, and $\xi_i \varepsilon \mathcal{R}$, $i = 1, 2, 3, \ldots,$ and $\xi_i \to \xi_1$ as $i \to \infty$, so we must have $\xi_1 \varepsilon \mathcal{R}$. Whence there is a $(\phi,u)\varepsilon H \times \mathcal{U}(\Omega)$, where u has domain $[\omega,t_2]$ (t_2 not necessarily equal to t_1), such that $x(t_2,\phi,u) = \xi_1$. It can, however, be shown in the same way as was done in the proof of Theorem 5.2.1 that $\xi_1 \varepsilon \mathcal{R}_{t_1}$. Consequently there is a $(\phi,u)\varepsilon H \times \mathcal{U}(\Omega)$, u has domain $[\omega,t_1]$ such that $x(t_1,\phi,u) = \xi_1$. Whence $\xi_1 = x(t_1,\phi,u)\varepsilon F(t_1)$, and $(\phi,u)\varepsilon \hat{\mathcal{U}}(H,\Omega)$, thereby showing that Q is closed, and therefore compact.

Let $\Pi : R^n \to R :: x = (x^1,\ldots,x^n) \to \Pi(x) = x^1$, then the mapping Π is continuous. Since Q is compact, it follows that $\Pi(Q)$ is a compact subset of R. Also $Q \neq \emptyset$ (by hypothesis) so the conclusion of Theorem (5.2.2) now follows immediately.

BIBLIOGRAPHY

Books

I. Alexandroff, P. and Hopf, H. Topologie. erster Band, Verlag von J. Springer in Berlin, 1935.

II. Bateman, P. T., Radstrom, H., Hanner, O., Macbeath, A. M., Rogers, C. A., and Klee, V. L. Seminar on Convex Sets. The Institute for Advanced Study, Princeton, N. J., 1949-1950.

III. Bellman, R. and Cooke, K. L. Differential-Difference Equations. Academic Press, New York, 1963.

IV. Birkhoff, G. and Rota, Gian-Carlo. Ordinary Differential Equations. Ginn and Co., 1962.

V. Bourbaki, N. Topologie Générale. Hermann, Paris, 1940-1949.

VI. Coddington, E. A. and Levinson, N. Theory of Ordinary Differential Equations. McGraw-Hill, New York, 1955.

VII. Eggleston, H. G. Convexity. Cambridge at the University Press, 1958.

VIII. Hausdorff, F. Set Theory. translated from the German by Aumann, John R., et al, second edition, Chelsea, New York, 1962.

IX. Kelley, J. L. General Topology. Van Nostrand, New York, 1955.

X. Kolmogorov, A. N. and Fomin, V. J. Elements of Functional Analysis. Vol. 1, Graylock Press, Albany, N. Y., 1957.

XI. Kolmogorov, A. N. and Fomin, V. J. Elements of Functional Analysis. Vol. 2, Graylock Press, Albany, N. Y., 1961.

XII. Liusternik, L. A. and Sobolev, V. J. Elements of Functional Analysis. translated from the Russian by A. E. Labarre, et al, F. Ungar, New York, 1961.

XIII. McShane, E. J. Integration. Princeton University Press, 1944.

XIV. McShane, E. J. and Botts, T. A. Real Analysis. Van Nostrand, Princeton, N. J., 1959.

XV. Pontryagin, L. S., Boltyanskii, V. J., Gamkrelidze, R. V., and Mischenko, E. F. The Mathematical Theory of Optimal Processes. translated from the Russian by K. N. Trirogoff, edited by L. W. Neustadt, John Wiley and Sons, 1962.

XVI. Riesz, F. and Sz-Nagy, B. Functional Analysis. translated from the second French edition by L. F. Boron, F. Ungar, New York, 1955.

XVII. Struble, R. A. Nonlinear Differential Equations. McGraw-Hill, New York, 1962.

Articles

1. Balakrishnan, A. B. "An Operator Theoretic Formulation of a Class of Control Problems and a Steepest Descent Method of Solution," SIAM Journal, Control, Series A, I, (1963), pp. 109-127.

2. Blackwell, D. "The Range of Certain Vector Integrals," Proceedings of the American Mathematical Society, 2 (1951), pp. 390-395.

3. Chang, Albert, "An Optimal Regulator Problem," SIAM Journal, Control, Series A, II (1965), pp. 220-233.

4. Choquet, G. "Convergences," Annales de l'Université de Grenoble, vol. 23 (1947-1948), pp. 55-112.

5. Datko, R. "An Implicit Function Theorem with an Application to Control Theory," Michigan Journal of Mathematics, 11 (1964), pp. 345-351.

6. Ewing, G. M. "Sufficient Conditions for Global Extrema in the Calculus of Variations," to be published: Journal of the Astronautical Sciences.

7. Filippov, A. F. "On Certain Questions in the Theory of Optimal Control," SIAM Journal, Control, Series A, I (1963), pp. 76-84

8. Hestenes, M. R., "Applications of the Theory of Quadratic Forms in Hilbert Space to the Calculus of Variations," Pacific Journal of Mathematics, 1 (1951), pp. 525-580.

9. Kuratowski, C. "Les Fonctions Semi-Continues dans l'Espace des ensembles fermés," Fund. Math., 18 (1932) pp. 148-160.

10. LaSalle, J. P. "The Time Optimal Control Problem," Contributions to the Theory of Nonlinear Oscillations, vol. V, pp. 1-24, Princeton University Press, Princeton, N. J., 1960.

11. Lee, E. B. and Markus, L. "Optimal Control for Nonlinear Processes," Archive for Rational Mechanics and Analysis, 8 (1961), pp. 36-58.

12. Michael, E. "Topologies on Spaces of Subsets," Transactions of the American Mathematical Society, 71 (195.), pp. 152-182.

13. Neustadt, L. W. "The Existence of Optimal Controls in the Absence of Convexity Conditions," Journal of Mathematical Analysis and Applications, 7 (1963), pp. 110-117.

14. Oǧuztörelli, M. N. "A Time Optimal Control Problem for Systems Described by Differential Difference Equations," SIAM Journal, Control, Series A, I (1963), pp. 290-310.

15. Oǧuztörelli, M. N. "Relay Type Control Systems with Retardation and Switching Delay," SIAM Journal, Control, Series A, I (1963), pp. 275-289.

16. Robertson, A. P. and Wendy. "On the Space of Subsets of a Uniform Space," Proceedings of the American Mathematical Society, 12 (1961), pp. 321-326.

17. Roxin, E. "The Existence of Optimal Controls," Michigan Mathematical Journal, 9 (1962), pp. 109-119.

18. Warga, J. "Relaxed Variational Problems," Journal of Mathematical Analysis and Applications, 4 (1962), pp. 111-128.